Summer 4-26-2013

Child Maltreatment Prevention and Health Promotion: Examining the Effectiveness of a Nurse Home-Visiting Program

Paul Lanier
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Brown School of Social Work

Dissertation Examination Committee:
Melissa Jonson-Reid, Chair
John Constantino
Brett Drake
Patricia Kohl
Doug Luke
Nancy Weaver

Child Maltreatment Prevention and Health Promotion: Examining the Effectiveness of a Nurse Home-Visiting Program

by

Paul Jetter Lanier III

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2013

St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgements

This dissertation represents the tremendous support from world-class professional mentors and a dedicated personal team of friends and family. I am particularly grateful to Melissa Jonson-Reid for shepherding me through this process and challenging me to always improve my work. I appreciate Brett Drake and Trish Kohl for their mentorship that prepared me for this dissertation and for their constant support of my career goals. They are amazing coaches and cheerleaders. I thank John Constantino for pushing me to make the most impact possible with this opportunity. Thanks to Doug Luke for championing the cause of methodological rigor and Nancy Weaver for keeping my work grounded in the public health goals of the program. This committee was truly my dissertation "dream team", supporting and challenging me from start to finish.

I would also like to acknowledge the funders of my doctoral education and dissertation research. I was fortunate to receive an NIMH pre-doctoral fellowship (T32MH019960), Center for Violence and Injury Prevention PhD Scholar award funded by the CDC (R49CE00150-03), and a Doris Duke Foundation Fellowship for the Promotion of Child Well-Being. I thank the funders and Enola Proctor, Melissa Jonson-Reid, and Deborah Daro for investing your time in supporting doctoral education and research.

This dissertation would certainly not have been possible without the leadership of Nurses for Newborns. Melinda Ohlemiller accepted me as part of the NFN family and agreed to serve as my policy mentor. Angie Recktenwald was by my side in the trenches for months. Their patience and persistence is a testament to their passion for improving the work that they do. I also want to acknowledge the thousands of families who participated in NFN and also agreed to support research perhaps in hopes of improving the quality of services for families to come.

I am lucky to have such a strong network of personal support. Special thanks to my dad, Libby, Shannon, Jeff, and my old North Carolina and new St. Louis friends. Sharon, as usual, you exemplified unending commitment to helping others. To my family, Virginia, Jett, and Cal, you are my sunshine. Thank you for being so supportive of me. This is for you.

ABSTRACT OF THE DISSERTATION

Child Maltreatment Prevention and Health Promotion:

Examining the Effectiveness of a Nurse Home-Visiting Program

by

Paul Jetter Lanier III

Doctor of Philosophy in Social Work

Washington University in St. Louis, 2013

Professor Melissa Jonson-Reid, Chair

Problem: Child abuse and neglect is a devastating, yet preventable, social problem. Early childhood home visiting services are currently considered the most promising approach to maltreatment prevention. Expansion of evidence-based home visiting services authorized by the Affordable Care Act has created a need to better understand the utilization and effectiveness of preventive services delivered in the community. Nurses for Newborns is a nurse home visiting program serving high-risk caregivers and medically-fragile newborns using a flexible, client-driven service model.

Methods: This study used longitudinal services data linked with administrative child welfare data. A sample of low-income families (n = 3,620) who received services from 2009 to 2011 were followed through the end of 2012. Analyses first focused on describing the service population and identifying predictors of engagement and retention. Next, child maltreatment, child development, and maternal mental health outcomes were compared across policy-relevant subgroups. Last, a quasi-experimental design using propensity score analytic methods was conducted to identify a causal treatment effect for maltreatment prevention.

Results: The program serves a very high-risk population with individuals experiencing multiple social stressors in addition to medical issues relating to the pregnancy or newborn. There is variation in the level of service use, but generally higher risk families are more likely to engage and participate in services longer. Subgroups of families were found to have higher rates of child developmental concerns, caregiver stress, postpartum depression, and later maltreatment. Families who enroll prenatally appear to be a very different group in terms of risk factors and maltreatment outcomes. Only 1% of families who enroll prenatally have a later maltreatment report compared to 19% of those who begin postpartum. Among postpartum families, propensity score matching was successful in balancing an engaged treatment group with a dropout comparison group. There was not a significant difference between these groups in risk for later maltreatment report. Among families with a maltreatment report prior to home visiting services, the risk of maltreatment was significantly lower for those in the treatment group.

Conclusion: Consistent with prior home visiting research, this study found troubling levels of attrition and was not able to detect a statistically significant difference in overall risk of later maltreatment report. Families who receive services prenatally have an exceptionally low rate of later report despite high levels of risk. The program was successful in preventing maltreatment recurrence among those families with a prior report. While some programs serve only first-time mothers, this study found that multiparous clients had some of the highest levels of risk and the poorest outcomes. Lack of overall program findings may be at least partially attributable to low service dosage. This study indicates that NFN is making a measureable impact in maltreatment prevention among certain subgroups of families.

Chapter 1: Overview and Research Aims

Introduction

The experience of child abuse and neglect, especially in the early years of development, has a profound impact on the health and well-being of child victims. It can greatly limit opportunities for healthy and productive childhood years and permanently alter an individual's well-being through the life course (Shonkoff, Boyce, & McEwen, 2009). Child abuse and neglect (maltreatment) occurs within a "toxic relational environment that poses considerable risk for maladaptation across biological and psychological domains of development" (Cicchetti & Toth, 2005, p. 410). Given the high rates of maltreatment, especially among vulnerable populations, and the adverse outcomes associated with a history of maltreatment, there is a clear moral and public health imperative to prevent maltreatment from ever occurring. Evidence-based early childhood home visiting has become the primary approach to prevent maltreatment and support families with young children. The current policy emphasis on program expansion at the federal and state level demands rigorous evaluation of existing services and dissemination of new knowledge.

A strong rationale for addressing maltreatment as a public health problem exists due to the numerous negative health and mental health sequelae that have been linked with maltreatment in childhood. However, our knowledge about these negative outcomes far outweighs what we know about effective preventive strategies (MacMillan, 2010). Furthermore, the existing child protection system has historically focused on responding to maltreatment after it occurs. The investment in prevention is usually a small percentage of allocated resources and prevention programs are often the first to be cut when balancing budgets (Thomas, Leicht, Hughes, Madigan, & Dowell, 2003). Most current prevention programs are understudied, underfunded, and may not use what is known about evidence-based practice and policy. One

approach to addressing this problem is to shift priorities towards improving research on existing prevention programs. This is essential to understanding how home visits are being delivered in the field and to inform the implementation of evidence-based approaches by improving the fit of services with client needs.

Although varied in delivery and content, home-visiting programs focus on building supportive relationships with caregivers in their home environment to improve maternal and child health outcomes. Several models have had significant research attention including randomized controlled trials with favorable results for certain populations (Geeraert, Van den Noortgate, Grietens, & Onghena, 2004; Sweet & Applebaum, 2004). There are, however, many different forms that home visitation can take and overall the results have been mixed (Astuto & Allen, 2009; Boonstra, 2009). The Maternal, Infant, and Early Childhood Home Visiting Program, established under the Patient Protection and Affordable Care Act of 2010, provides $1.5 billion in mandatory funds over five years for states to implement and expand home-visiting models that are "grounded in empirically-based knowledge" and that have been rigorously evaluated. While an empirical and theoretical foundation has been constructed regarding the efficacy of home-visiting models in controlled studies with specific populations like first time mothers (Olds, 2008), much is still unknown about the effectiveness of home visitation when delivered in the community with higher risk families within their current service framework. (Thompson, Kropenske, Heinicke, Gomby, & Halfon, 2001). It is not clear that such highly-researched programs will be effective with other groups and across all desired domains. Nor is it clear that other less-researched models are lacking in effectiveness, especially among groups not previously studied.

Given the clear intent to support evidence-based practices and a national emphasis on improving outcomes for children in low-income families, policymakers have an urgent need for practice-based research that both helps move from "bench" to the "trench" as well as informs the "bench" with what is already happening in the "trench". This study helps fill this gap in knowledge by examining key components of a pre-existing home visitation program that specifically targets high risk mothers and infants to understand how aspects of this program like dose of nurse home visits impact child health and prevent maltreatment. Nurses for Newborns (NFN) is an established program that shares some commonalities with a well-research nurse home visiting program, Nurse Family Partnership (NFP, "Olds Model"), but also has several unique features that provide advantages to advance research that informs policy and practice. The next chapter will elaborate further regarding the similarities and differences between these models.

Statement of the Problem

The prevention of childhood maltreatment and promotion of child well-being is an investment in the future of the child, his or her family, and society. Intervening with families during a child's first months and years may provide the best way to keep families together, prepare caregivers for effective parenting, and promote healthy development. Early and recurrent maltreatment has been shown to disrupt brain development and damage regulatory systems essential for normal functioning (Shonkoff, Boyce, & McEwen, 2009). Early childhood maltreatment poses a serious risk to biological and psychosocial development warranting a significant public health priority to support prevention (Cicchetti & Toth, 2005; Whitaker, Lutzker, & Shelley, 2005).

Prevalence. According to the 20th annual federal publication using data from the National Child Abuse and Neglect Data System, there were an estimated 3.3 million child welfare referrals involving the alleged maltreatment of 6 million children in the United States during 2009 (US DHHS, 2010). Of these referrals, about 1.5 million (25%) were classified as victims of maltreatment following investigation. Children aged zero to one had the highest rate of victimization at 20.6 per 1,000 children. Neglect (78.3%) was by far the most common type of maltreatment reported, followed by physical abuse (17.8%), sexual abuse (9.5%), and psychological maltreatment (7.6%). There were also an estimated 1,770 child fatalities (2.3 per 100,000 children) that resulted from abuse or neglect.

In an attempt to estimate periodic counts of actual maltreatment the National Incidence Study (NIS) measures the prevalence of maltreatment beyond the children investigated by child protective services (Sedlak, et al., 2010). The NIS employs the federally legislated harm and endangerment definitional standards to determine whether or not a case "counts" as maltreatment (Sedlak, et al., 2010). The NIS-4 found an estimated 1,256,600 children in the United States (one out of 58) were victims of maltreatment under the more-stringent harm standard. Of cases that reached the harm standard, 44% were identified as abuse and 61% were neglect. An additional 2,905,800 children (one out of 25) experience maltreatment classified at the level of the endangerment standard. For these cases, 29% were abuse and 77% were neglect.

Child abuse and neglect occurs in the context of a dysfunctional home and society. It is therefore difficult to separate the effect of the experience of maltreatment from other confounding psychosocial correlates. However, research has found that many deleterious outcomes can be uniquely linked with the individual experience of child maltreatment. Studies have emerge over the past decades establishing a strong causal link between exposure to child

maltreatment and a range of maladaptive adult health behaviors and significant chronic health and mental health problems leading to a marked reduction in life expectancy (Kendall-Tackett, 2002; Corso, Edwards, Fang, & Mercy, 2008; Lanier, et al., 2010).

One scholar has suggested that "child maltreatment is the single most preventable and intervenable contributor to child and adult mental illness in the country" (DeBellis, 2001, p. 539). Other lines of research have examined the effect of maltreatment on a range of social consequences including intellectual and academic outcomes (Perez & Widom, 1994; Jonson-Reid, et al., 2004), juvenile and adult criminal offending (Jonson-Reid, 1998, 2002), interpersonal relationships (Colman & Widom, 2004), personality development (Johnson, et al., 1999; Kim, Cicchetti, Rogosch, & Manly, 2009), economic well-being (Currie & Widom, 2010), and future abusive and neglectful parenting behavior (Egeland, Jacobvitz, & Sroufe, 1988; Berlin, Appleyard, & Dodge, 2011). Given the impact of maltreatment on so many aspects of well-being, identifying effective strategies to prevent maltreatment has potentially far-reaching implications.

The burden of child abuse and neglect impacts low-income children at a higher rate than children in families of higher socioeconomic (SES) status (Sedlak, et al., 2010). Child poverty is consistently marked as a strong and consistent predictor of child welfare system contact and chronic maltreatment (Drake, Jonson-Reid, Way, & Chung, 2003; Loman, 2006; Jonson-Reid, Emery, Drake, & Stahlschmidt, 2010). Poor children often experience the cumulative burden of an impoverished home environment, harsh or neglectful parenting, and an under-resourced community environment (Duncan, Brooks-Gunn, & Klebanov, 1994; Brooks-Gunn & Duncan, 1997). Children in lower SES groups are more likely to have poor pediatric health trajectories beginning with higher risk for infant mortality (Singh & Kogan, 2007) and poor infant health

outcomes (Olson, Diekema, Elliott, & Renier, 2010) then higher rates of childhood injury (Chen, Matthews, & Boyce, 2002; Laflamme, Hasselberg, & Burrows, 2010) and developmental delay (Jonson-Reid, Drake, Kim, Porterfield, & Han, 2004; Hackman, Farah, & Meaney, 2010; Engle & Black, 2008). For example, the infant mortality rate for a child born to a mother with less than a high school diploma is 8.3 per 1,000 births compared to 3.8 for a mother with a bachelor's degree or higher (Matthew & MacDorman, 2010).

Child poverty is far from a rare event in the United States and has become the unfortunate norm for almost a true majority of infants and toddlers. Of the more than 11 million children between age zero and three in 2010, 48% live in poverty, up from 44% in 2005 (Addy & Wight, 2012). Further perpetuating racial disproportionality in key child outcomes (Drake, et al., 2011), the poverty rate for black children is consistently twice that of white children (70% vs. 35%). In examining areas for intervention for child maltreatment and health promotion, the powerful impact of poverty cannot be ignored.

While the risks of poverty and maltreatment often co-occur and exacerbate damaging effects, abuse and neglect is known to have a profound singular impact on child development. Research on neural plasticity suggests that early brain development depends upon interaction and stimulation from the environment, directly implicating infant abuse and neglect with neurobiological consequences (Cicchetti & Rogosch, 2001; Olson, Bates, Sandy & Schilling, 2002; DeBellis, 2007; Nemeroff, 2004; Teicher, Dumont, Ito, Vaituzis, Giedd & Anderson, 2004). Intellectual delays can occur from lack of appropriate stimulation while disruptions in emotional and cognitive processing systems are likely linked to neurochemical responses to the stress of maltreatment (Perry, 2000; American Academy of Pediatrics, 2001; Strathearn, Gray,

O'Callaghan & Wood, 2001). Early childhood is a fragile period of time where the groundwork for lifelong health and well-being is established.

Costs. Reliable estimates of the true cost of maltreatment to an individual and society are difficult to obtain (Courtney, 1999; Corso & Fertig, 2010). While the human service costs of investigating a report of maltreatment and the subsequent economic impact on a variety of public service systems can be quantified, the toll on the individual's well-being and quality of life is much more difficult to measure. There have been several attempts to conduct cost of illness (COI) analyses examining the direct and indirect costs associated with maltreatment in the United States (Corso & Lutzker, 2006). Prevent Child Abuse America has conducted two national COI analyses that are commonly cited as comprehensive estimates. In 2001, the estimated total cost was $94 billion (1993 US dollars, $135 billion in 2007 US dollars) with $24 billion in direct costs and $70 billion in indirect costs (Fromm, 2001). In 2007, the economic burden was estimated to be $104 billion, a 25% reduction from the previous estimate (Wang & Holton, 2007). Direct costs ($33 billion) included hospitalizations, mental health care, child welfare services, and law enforcement and indirect costs ($71 billion) included special education, juvenile delinquency, adult mental health and health care, lost productivity to society, and adult criminality. A recent review by Corso & Fertig (2010) has adjusted the 2007 estimate to $65 billion reflecting a methodological overestimation in the Wang & Holton calculation.

It is important to note, however, that the level of severity of maltreatment used to define the base number for cost calculation is about that of the NIS 'harm standard' or the 'victim' level of NCANDS data. There are many studies that indicate that children with "unsubstantiated" cases or cases determined to be less severe have about equal rates of poor outcomes (Drake et al,

2003; Hussey, et al., 2005). Given that about two-thirds of reports nationally are found to unsubstantiated (US DHHS, 2010), these cost studies are likely a gross underestimation.

While the actual total cost of maltreatment remains elusive, there are clearly large burdens to the individual and society that could be eliminated by preventing abuse and neglect. Many are beginning to acknowledge the power of parenting relationships in impacting the life course of the individual and the disproportionate impact on underserved populations. The quality of the family environment in the first years of life set in motion a trajectory that can lead to tragically different outcomes.

The high cost of ignoring the problem of maltreatment and failing to support at-risk caregivers is beginning to gain attention in fields such as education and economics. Heckman (2008) states that it is this "accident of birth", or the lack of choice of one's parents, that is the first constraint on all of an individual's later outcomes. Early interventions that shift the childhood experience away from early adversity and seek to create "safe, stable, and nurturing environments" can improve outcomes and potentially avoid lifelong costs associated with impairment (Mercy & Saul, 2009, p. 2262). Highlighting this return on investment in prevention is critical to the success of programs such as home visiting that seek to divert future costs. A recent book, "How Children Succeed" by Paul Tough (2012), linked the recent advances in developmental neuroscience to provide a strong rationale for focusing on the home environment during infancy in order to close the achievement gap in educational outcomes. The author suggests that the solution to ensuring that children, particularly those living in poverty, are psychologically prepared for school as well as later vocational success as adults may be rather simple:

> And how do you that? It is not magic. First, as much as possible, you protect him from serious trauma and chronic stress; then, even more important, you provide him with a secure, nurturing relationship with at least one parent and ideally two. (p. 182)

Evidence for Home Visiting

Home-visiting services have expanded to serve an estimated half million families in 40 states and are slowly moving towards universal delivery despite mixed evidence of effectiveness in preventing maltreatment (Astuto & Allen, 2009; Howard & Brooks-Gunn, 2009). While this is a promising expansion in services and a movement in the right direction of serving families that can benefit from services, there is likely still large unmet need for the 23 million children aged zero to five in the US and the 4 million new births each year (Daro, 2009). Further, if the expansion includes program components that are less efficacious there is a danger that the support for the services will erode.

The majority of evidence supporting home visiting for the prevention of maltreatment has come from a single model (Nurse-Family Partnership) and a specific target population including only first-time mothers enrolled prenatally (Daro, Dodge, Weiss, & Zigler, 2009). This model has shown promising results in a number of key domains across three randomized controlled trials (RCTs) with long follow-up periods, although evidence related to the prevention of child abuse is not definitive. Another widely disseminated model, Healthy Families America, was developed by Prevent Child Abuse America utilizing paraprofessional "Family Service Workers" to deliver home visiting services. In several RCTs, this model has failed to show significant reduction in rates of child abuse and neglect reports. However, controlled trials give only an indication of the efficacy of an intervention under ideal conditions.

There is a need to study the details of services provided within home visitation programs that operate in "real world" communities without carefully selected families. Further, within this

context, more investigation is needed to determine under what circumstances home visiting services can be successful in preventing maltreatment. Only then can we begin to understand what combinations of services work with what types of populations in what types of communities. Without this knowledge, we will continue to lack the empirical evidence needed to inform policy and program implementation on the widest scale. This study explores the depth and breadth of home-visiting effectiveness in one such program.

Description of Nurses for Newborns Program

The NFN program was initially developed to fill the gap in service need created when new mothers return home from the hospital after delivery. NFN was developed in 1991 with a mission to provide a safety-net for the families at highest risk for infant mortality and child maltreatment. The families served by NFN often lack the required help once they return home and are often isolated from outside support (Rohrbach, 1993). The combination of exhaustion and a rush to discharge from expensive hospital stays creates a situation where families require additional information and support once they have returned home. Postpartum discharge places the infant at risk for undetected medical problems such as jaundice, dehydration, sepsis, and heart defects. The many challenging aspects of caring for a newborn such as breastfeeding and safe sleep must be adequately reinforced beyond the first days of life. The primary focus of NFN when it was founded was to detect problems that may have emerged during the first few days after discharge and address these problems with appropriate referrals and parent education. NFN has also expanded services to the Nashville, Tennessee area. This study will focus on those families served in Missouri only.

The initial target population for NFN was medically fragile newborns following a stay in the neonatal intensive care unit (NICU). This program provided in-home nursing and education

services for families. NFN currently operates programs that focus on serving four subpopulations: *Bridge to the Future* continues to serve families with medically fragile infants, *Safe Beginnings* serves parents with intellectual or physical disabilities or with a diagnosed mental illness, *Teen Parent* serves mothers under the age of 19, and *Bright Futures* serves mothers whose primary risk factor is poverty. While the service structure and delivery is consistent for all families, these four programs reflect the breadth of high-risk groups served by NFN. The sample for this dissertation will include families served across these four populations.

Key Research Aims

Currently, the most rigorous nurse home visiting studies come from heavily controlled intervention trials requiring prenatal enrollment and excluding mothers who have previous children, significant mental health disorders, or infants with serious medical conditions (MacMillan, 2009). Yet these are some of the populations who are likely most in need of parenting support. Further, while beginning services in the prenatal period or at birth may be ideal, recruitment methods for many programs may miss mothers who avoid or cannot access prenatal care or are unlikely to follow through on information provided at the hospital. Additionally, with changes in family and community dynamics, there is no guarantee that services for a first birth will offset increases in risk that may ensue with a subsequent child.

The purpose of this research is to examine what home visitation program characteristics are linked to child maltreatment prevention (official reports of abuse and neglect) among high risk and diverse families served by the Nurses for Newborns (NFN). This innovative research plan leverages unique features of NFN to address gaps in the evidence base to inform policy and practice. NFN is a well-established program operating in both urban and rural areas and has almost no exclusion criteria. Families are only denied services after a referral if the caregiver is

already receiving home visiting services from another agency or if the nurse determines there is no need for services at the first visit. NFN serves families who may have multiple children, families initiating services both prenatally and postnatally, mothers with disabilities, significant mental health, or substance abuse problems, and those with medically-fragile newborns. This feature allows the opportunity to explore outcomes for families across groups that may have not been studied in prior research. Understanding what outcomes can be expected for different groups of families is critical for informing policy and practice.

The impact of intervention dose on outcomes has lacked sufficient empirical attention. Many home visiting services are open-ended with number of visits limited only by the age of the child. Dosage information is often not provided and most studies provide only an average count of visits received or average hours to estimate dose in the same way that one would with a fixed-length intervention (Sweet & Applebaum, 2004). A problem with this approach is that it is impossible to discern the reason a family received fewer services than another. NFN electronic data includes detailed qualitative nurse comments for each visit. This qualitative data can potentially address a large portion of unexplained variance by exploring contextual influences that may influence engagement with services.

Like other voluntary, long-term interventions, early dropout from home visiting services is one of "great challenges facing home visitation" (Ammerman, et al., 2006) with refusal rates ranging from 8 to 22% (McCurdy, Daro, et al., 2006) and attrition rates ranging from 20 to 67% and (Damashek, Doughty, Ware, & Silovsky, 2011). As home visiting services are scaled up in community settings, more research is needed to determine the barriers to service engagement and retention. This study explores the impact of the family, nurse, and community in predicting these engagement outcomes.

In addressing these issues, the aims of the proposed research are:

Aim 1. To describe service use patterns and explore whether dose moderates outcomes.

Question 1.1: What proportion of families receives a second visit (initial engagement) and what is the distribution of number of visits, hours of service contact, and weeks of agency contact (retention)?

Question 1.2: To what extent do engagement and retention vary by individual, provider, and neighborhood?

Question 1.3: Which individual-level predictors are associated with engagement and retention?

Question 1.4: Is retention in services/dosage associated with skills training and documented improvements in parenting knowledge?

Question 1.5: Is retention in services/dosage associated with decreased risk for maltreatment?

Aim 2. To compare maltreatment reports, child development, and maternal mental health outcomes among those served in key policy-relevant subgroups including:

- first-time mothers and those with previous births (multiparous vs. primiparous)
- prenatal and postnatal referrals
- healthy and medically-fragile newborns
- mothers with substance abuse/mental health problems and those without
- teen mothers and older mothers
- urban and rural families
- race subgroups

Question 2.1: Are there differences in rates of child maltreatment reports for family subgroups?

Question 2.2: Are there differences in maternal mental health indicators for family subgroups?

Question 2.3: Are there differences in child developmental health indicators for family subgroups?

Aim 3. To compare subsequent child maltreatment reports for families that receive NFN services to a matched comparison group using a quasi-experimental design utilizing propensity score and survival analysis.

Question 3.1: To what extent can any differences in the *rate* of maltreatment reports be attributed to the Nurses for Newborns intervention?

Question 3.2: To what extent can any differences in the *timing* of maltreatment reports be attributed to the Nurses for Newborns intervention?

Chapter 2: Theoretical and Empirical Background

This chapter will introduce the theoretical perspectives considering the etiology and prevention of child maltreatment with a focus on the theoretical context of home visiting interventions. The first section will briefly review the evolution of child abuse and neglect as an identified social problem and introduce the theories that attempt to explain the causes of child maltreatment. Then the discussion will shift from description of the problem of maltreatment, to the public health prevention framework. Within this model, theory is used to inform specific prevention approaches, particularly home visiting.

Following the discussion of theory, this chapter will also review the key studies that have evaluated the effectiveness of home visiting programs to prevent maltreatment, for which there is a relatively strong base of empirical evidence. Appendix A provides a description of the major peer-reviewed articles that have assessed child maltreatment outcomes in home visiting evaluations. Appendix B provides a list of review articles, systematic reviews, and meta-analyses that have attempted to summarize the results of home visiting research.

Theoretical Context of Child Maltreatment

Maltreatment is a complex phenomenon and the theories relating to the etiology of child maltreatment have widely varied, exploring factors across expanding levels of human ecology. Understanding the causes of maltreatment directly informs the theoretical base for the prevention of maltreatment. Our reactive response, out of the urgent need to protect children, has preceded a clear understanding of how to prevent the problem from occurring in the first place. To prevent maltreatment, it is crucial to first understand why maltreatment occurs.

Maltreatment as a social problem. In the United States, the 1874 case of "Mary Ellen" first exemplified the extent of child maltreatment in our society and the lack of attention the

problem historically received (Cicchetti & Carlson, 1989). A New York social worker found this 8-year-old girl chained, beaten, and starved by her adoptive parents but had no formal child protective service agency to report the abuse. The New York Police Department refused the case due to inexistence of laws protecting children from their parents. Instead of a formal system, the founder of the New York Society for the Prevention of Cruelty to Animals individually intervened and finally brought the case to trial. The publicity of the case and the subsequent one year sentence of the adoptive mother led to the founding of the Society for the Prevention of Cruelty to Children in 1875. This new-found interest in child protection led to formation of agencies, hospitals, and homes for abandoned or abused children (Cicchetti & Carlson, 1989). Later, thinking would expand to include not only strategies to intervene in active cases of abuse but also in identify those families at risk and prevent maltreatment from ever occurring.

Almost 100 years later, the "battered child syndrome" was first introduced by Kempe and his colleagues (Kempe, Silverman, Steele, Droegemuller, & Silver, 1962) sparking the beginning of modern theory development of maltreatment etiology (Newberger & Newberger, 1982). This new phenomenon was identified primarily by the confluence of radiological data of skeletal imaging that suggested trends in injuries found in children represented a new medical classification or "syndrome". Within a few short years, federal and state policy regarding child protection and welfare had been created.

Etiology of maltreating behavior. With public awareness piqued, many questions remained. Prime amongst these was, why do parents and caregivers harm their children? The focus of the early research on battered-child syndrome was to profile caregivers and determine the psychopathological dysfunctions that lead to abuse and neglect (National Research Council, 1993). This individual-based approach was criticized for its neglect of powerful community-

level pressures that may impact family functioning. So, more sociological approaches were posited to include the social, political, and cultural environment as factors influencing the risk of maltreatment (Gil, 1970; Garbarino, 1977). Over time, researchers have found that in order to explain the heterogeneity in precursors and outcomes of child maltreatment, a broader theory that incorporated dyadic, familial, and environmental variables is necessary (Belsky, 1980; National Research Council, 1993). In examining the causes of child maltreatment and areas for intervention, theorists generally apply some version of an ecological framework of human development (Bronfenbrenner, 1979) to describe interactions of risk and protective factors across multiple overlapping contexts (Widom, 2000; Sidebotham, 2001; Bruskas, 2008; Zielinski & Bradshaw, 2006). Ecological theory is discussed in more detail later.

Public Health Framework for Prevention

Mounting epidemiologic evidence outlining the impact of child abuse and neglect on individuals and society has provided a strong impetus to consider maltreatment a public health problem. There is currently much agreement that shifting towards a public health approach is necessary for maltreatment prevention to be effective in terms of a lens for research and service delivery (Stagner & Lansing, 2009; Prinz, Sanders, Shapiro, Whitaker, & Lutzker, 2009). This public health approach is the cornerstone of the current maltreatment prevention strategy.

The Director of the Division of Violence Prevention within the National Center for Injury Prevention at the Centers for Disease Control and Prevention (CDC) laid the groundwork in a paper outlining the role of the CDC and a broader public health approach to maltreatment prevention (Hammond, 2003). This paper delineates the difference between a public health approach and the existing criminal justice and child protection systems suggesting that while the approach is different, it can be complementary to existing systems. He also suggests that the

field follow other injury and violence prevention researchers in adopting the public health model by applying four steps to their approach to maltreatment prevention (Doll, Saul, & Elder, 2007). This framework begins by first utilizing epidemiological surveillance data to understand the scope of the problem. Second, the causes of maltreatment are identified by exploring risk and protective factors. Third, programs and policies that target child maltreatment prevention are evaluated. Finally, a public health approach encourages widespread adoption of the approaches that have strong empirical support.

Preventive interventions are defined as "strategies or a series of strategies that are implemented with the goal of preventing, reducing, or ameliorating injuries" (Doll, Saul, & Elder, 2007, p. 22). Previously, preventive interventions were categorized using terms from clinical medicine and disease prevention as either primary (preventing new cases), secondary (reducing established cases), or tertiary (decreasing disability in cases) levels of prevention. Gordon (1983) developed a classification system based on population risk levels that is more applicable to maltreatment prevention and more commonly used among public health researchers. This system targets populations based on level of risk and includes universal (general population), selective (populations with increased risk), and indicated (populations already exposed or with above average risk) preventive measures. Within this rubric, universal prevention would be provided to all families in the general population, selective interventions may target families with somewhat higher risk such as first-time parents, and indicated interventions would engage families already reported for maltreatment or where symptoms are beginning to emerge. Universal and selective approaches can be seen as proactive or "before-the fact" and indicated approaches are more reactive or "after-the-fact" (Stagner & Lansing, 2009).

With regard to maltreatment prevention, the goal of the CDC is "to create a social context

in which child maltreatment is not tolerated, and in which prevention and intervention services are evidence-based, effective, widely available, and socially valued" (Hammond, 2003, p.83). Within each of the steps of the public health approach there are concrete ways to achieve this goal. There is an opportunity to begin this change and improve the current system by expanding surveillance systems and identifying a better-defined range of risk factors. This will take a concerted and organized effort from state and federal policymakers, researchers, and other stakeholders to invest in improving the surveillance infrastructure to expand our understanding of the scope of maltreatment nationally and locally.

It could be argued that while the current surveillance system and knowledge regarding risk and protective factors are not perfect, these first two pieces of the public health model have been adequately developed to move forward with testing and disseminating interventions. A recent policy statement (Garner, Shonkoff, Siegel, et al., 2012) and technical report (Shonkoff & Garner, 2012) from the American Academy of Pediatrics outlined the need to further advocate for the development and implementation of effective prevention programs to address this growing public health imperative. The prevention of maltreatment is a priority of the CDC (Whitaker, Lutzker, & Shelley, 2005) and is a national objective for Healthy People 2020 (US DHHS, 2012). The discussion will now turn to the theoretical development of specific maltreatment prevention strategies.

Theoretical Basis for Maltreatment Prevention Strategies

This section will describe the application of theory to the development of preventive interventions, specifically home visiting. Much of the early focus of prevention centered on protection and punitive responses to identified maltreatment and to some extent on preventing maltreatment recurrence through rehabilitative services (Stagner & Lansing, 2009). The goal of

these efforts was to develop a public infrastructure that could respond to cases of maltreatment and remove the child from the reach of the maltreating caregiver to disrupt the abusive or neglectful behavior by placing the child in an alternative living situation that was assumed to be safe. Future maltreatment was thereby prevented by physically removing the child from the source of the behavior. Laws were also created that criminalized abuse or neglect. This was assumed to have a deterrent effect, particularly in preventing cases of sexual abuse (Finkelhor, 2009). This approach was clearly not a sufficient response to prevent new cases of abuse and neglect.

The reactive child welfare system focuses very little on identifying high-risk families and intervening before maltreatment occurs and even less on universal primary prevention for the general population. While developing a system to respond to maltreatment once it occurs was a critical and complicated undertaking, preventing maltreatment requires a sophisticated understanding of the etiology of maltreatment that can inform effective change of risky interaction patterns. This shift towards early prevention and intervention is now a major focus for many in the child development and child welfare community. There are many examples of how theory is being used to inform these prevention practices.

Protective factors framework. A starting point for identifying a theoretical basis for prevention practice is to identify the core protective factors that reduce the risk for maltreatment (Stagner & Lansing, 2009). Based on empirical literature, five protective factors that decrease the risk for maltreatment have been identified by the Children's Bureau and are being infused in many prevention models. Attachment and nurturing with family members, knowledge of parenting and child development, parent emotional resilience, social connections and emotional support, and concrete supports with basic resources are known to play an interconnected role in

reducing parenting stress and improving positive parenting skills (Child Welfare Information Gateway, 2008). Home visiting programs, and other preventive interventions, often seek to develop or strengthen these specific protective factors.

The protective factors that were selected for focus were drawn from prior research and theory. Informing the five protective factors are four theoretical perspectives categorized by Daro (1993) based on earlier work by Newberger & Newberger (1982) that have been applied to maltreatment prevention. First, psychodynamic theory targets the maltreating caregiver and suggests that a person-oriented treatment plan that provides a better understanding and acceptance of their role as parents, the caregiver will be less likely to maltreatment their child (Steele & Pollack, 1974). Second, learning theory suggests that maltreating behavior is principally a function of lack of knowledge about effective and safe parenting strategies (Parke & Collmer, 1975). With more information about child development and parenting skills, caregivers will rely on less abusive and neglectful behavior.

Third, environmental theory addresses the protective factors of resources (Gil, 1970). This theory would suggest that with more abundant and better quality resources, abuse will be less likely. This theory acknowledges that parenting stress is determined by factors other than the parent-child relationship and that demands placed on the caregiver combined with a lack of support increase parenting stress. Fourth, ecological theory incorporates our understanding of the importance of social support and other community factors (Garbarino, 1977; Cicchetti & Lynch, 1993). With an increased network of community support and a positive community context, the family and environmental risk factors can be overcome leading to lower risk for maltreatment. Most prevention programs currently focus on some combination or variation of these protective factors. "Evidence-based" home visiting programs are certainly no exception, as

they have consistently described the enhancement of these types of protective factors as a primary focus of the intervention. The next section will describe in more detail the theoretical foundation for one specific home-visiting model.

Theories Informing Home Visiting

Perhaps the most comprehensively articulated theoretical framework for maltreatment prevention via home visiting has been in the development of the Nurse-Family Partnership home visitation model (Olds, 2006). Although Nurses for Newborns does not have a uniquely expressed theoretical model, the clinical guidelines and logic model reflect a very similar theoretical frame as has been proposed by NFP developers. While the importance of the human ecological perspective (Bronfenbrenner, 1979) was part of the early model, the mid-range theories of self-efficacy (Bandura, 1977) and attachment (Bowlby, 1969) are a focus of the home visitation theoretical model and further informed program development. Bronfenbrenner's (1992; 1995) *person-process-context-time* framework expands the focus of the ecological framework to explore the role of persons and processes, and not just social context, on human development. The *person* and *process* aspect of this framework can be further elaborated and applied to the home visitation model through the theories of attachment and self-efficacy. These three theories, ecological theory, self-efficacy theory, and attachment theory will be described further.

Ecological Perspective. The ecological perspective of child development (Bronfenbrenner, 1977; Bronfenbrenner, 1979) has been directly applied to the understanding of the etiology of maltreatment (Belsky, 1980; Belsky, 1993). Under this model, the individual exists within social systems arranged in expanding layers. In order to truly understand the context in which maltreatment exists, an analysis of the factors both proximal and distal to the

family must be included. The etiology of maltreatment is circular rather than linear with a combination of individual history, stress-producing forces, and the values of the society impacting the caregiver and family environment in which maltreatment may occur (Tzeng, Jackson, & Karlson, 1991).

Most current researchers consider the ecological framework the most valid way to truly examine maltreatment in research or to intervene with families at risk for maltreatment (National Academy of Sciences, 1993; Bruskas, 2008; Sidebotham, 2001; Zielinski & Bradshaw, 2006). This theory considers how individual, family, societal, and community factors combine and interact to create a situation where maltreatment would exist. This is one of the main advantages of this theory. The other advantage and perhaps its limitation is its acknowledgement of the complexity of maltreatment (Tzeng, Jackson, & Karlson, 1991). A key limitation is that the theory stops short of making any claims of causality or positing any testable hypotheses. This has led to various modifications and integration of other perspectives.

In a slight departure from the ecological model, Widom (2000) posited a "modified" ecological model to conceptualize the causes and consequences of childhood victimization. This model shifts the concentric circles of the Bronfenbrenner model to one that considers some areas of levels of the ecology overlapping and some areas not. This model reflects the fact that a child is not wholly enveloped by their experience within their family and over time, influence of peers and the community will likely increase. Also, parts of the child's experience may not involve the family or expanding levels of ecology. For example, a neglected child may have little exposure to community relationships and abused children may learn to dissociate from their home environment completely (Widom, 2000). Similarly, the family experience may not lie completely within the community environment and socially isolated families may by influenced very little

by the outside world. Despite these subtle differences, the importance of the environment, developmental change, and risk and protective factors at all levels of ecology is vital to the understanding of maltreatment.

Cicchetti & Lynch's ecological-transactional model (1993) extends the general ecological perspective by examining the interplay of risk and protective factors present across the social levels. The likelihood that maltreatment will occur is determined by a combination of factors that are present at the child, family, and societal level. This models seeks to explain the combined influence of child maltreatment and community violence on children's development over time and incorporates four nested levels of influence organized by proximity to the child. The "macrosystem" includes the societal factors and cultural beliefs that impact family functioning, the "exosystem" includes the neighborhood and community settings in which the child and family live and interact, the "microsystem" encompasses the immediate family environment, and the "ontogenic development" level incorporates the individual, reflecting the importance of the individual as an element within his or her own environment (Lynch & Cicchetti, 1998). This model is inherently complex and may reflect the true nature of the interactive relationship between an individual and the environment.

One purpose of the ecological-transactional model is to better define not only the causes of maltreatment but also the impact of this adversity on child development. Each level of the environment is thought to produce "potentiating" factors (increase the probability of poor outcomes) and "compensatory" factors (decrease the risk of poor outcomes) (Cicchetti & Rizley, 1981). Applying the temporal dimension to the model, potentiating and compensatory factors can be seen as either "transient" (temporary, fluctuating) or "enduring" (permanent conditions). Thus, risk factors that are enduring and proximal will have a greater impact on the individual

than those that are transient and distal. When buffering strong proximal potentiating factors, compensatory factors that are enduring will be more impactful than buffers that are transient (Lynch & Cicchetti, 1998).

This modified framework is quite useful in understanding the impact of maltreatment on health sequelae and in developing strategies for preventing maltreatment. Furthering our understanding of the interaction between risk/protective factors at the family level and risk/protective factors at the community level may provide new insights into approaches to conceptualize prevention methods that are tailored to the specific risk and protective profile of each child's ecological system. Simply focusing on the family unit as the target of intervention may miss opportunities for effective prevention strategies.

Attachment Theory. When studying human behaviors such as child maltreatment, behavioral ethological theorists propose the notion that although important, genetics and biology have a much smaller effect on behavior than learning and social environments (Browne & Parr, 1980). Studies of animal behavior have allowed researchers to manipulate and observe changes in parenting behavior when altering factors in the environment, leading to the development of attachment theory. The work of Harlow and colleagues (1963) found that infant rhesus monkeys that were socially isolated had trouble mating naturally when they became adults. In the laboratory setting, when the socially-isolated mothers reproduced they showed no maternal behavior, often attacking their offspring. Replicated studies of social isolation and lack of maternal attachment have found that "motherless mothers" abused and neglected their offspring with a dose effect of isolation (Browne & Parr, 1980). In another study, infant rhesus monkeys were presented with stressful stimuli without a real mother present. The infants had a choice between a terrycloth-covered artificial mother or a wire-frame with a milk bottle and teat.

Overwhelmingly, the infants chose to seek security with the terrycloth mother and not the wire-frame with the milk resource. These studies suggest that primates have developed an innate preference for maternal closeness for survival.

Bowlby's research attempted to connect the work of ethologists with psychoanalytic concepts in his development of attachment theory (Bowlby, 1969). From studies of mother-child bonding, Bowlby and his colleague Mary Ainsworth, explained the relationships that are formed during early childhood and how these relationships impact personality and mental health (Ainsworth & Bowlby, 1991; Bretherton, 1992). The basic principle of attachment theory is that an instinctual need for security and bonding has evolved among human infants and that sensitive, responsive, and accepting parenting is necessary for a healthy parent-child bond and for the healthy development of the child. Attachment theory primarily describes these early relationships as either secure or anxious. Anxious attachments, those that are insensitive, unresponsive, and unaccepting of the child's needs, can lead to maltreating behaviors.

Out of attachment theory, Ainsworth's work with the Strange Situation paradigm has been a common method for assessing the quality of parent-child attachment (Cicchetti & Carlson, 1989). In this structured laboratory technique, three stressors (introduction to a strange place, interaction with a new person, and brief separation from the mother) are initiated as common non-traumatic experiences that would illicit activation of attachment systems in infants (Ainsworth, Blehar, Waters, & Wall, 1978). The findings from numerous studies using this method have confirmed Bowlby's predictions by demonstrating that maltreated children form and measurably exhibit insecure attachment behaviors. Further studies led to the ability to observe and catalogue a mother's interaction style and successfully predict a child's reaction to the stress of the Strange Situation (Cicchetti & Carlson, 1989). Attachment theory is capable of

examining the parent-child relationship and characterizes the likelihood that a pattern of parenting consistent with maltreatment has developed or is likely to occur without intervention. Thus, attachment can serve as a marker for risk and also be a point of emphasis in promoting protective factors.

Although comprehensive and rooted in empirically vigorous fields of ethology and developmental psychology, attachment theory has limitations in its applicability to understanding the causes and outcomes of maltreatment. These limitations include: ignoring the importance of developmental needs beyond infancy, diminishing the significance of interactions with other children, adults, and institutions, and overemphasizing the role of parental impact on the development of undesirable outcomes (Tzeng, Jackson, & Karlson, 1991). This theory may describe the conditions and impacts of maltreatment in infancy based on the parent-child interactions, but fails to consider other developmental periods and factors beyond the interpersonal level. Although this theory does not speak of child abuse or maltreatment directly, an important contribution is its break with the psychoanalytic perspective that dominated the research and theory development of the time (Cicchetti & Carlson, 1989). Freudian emphasis on inner, unconscious dynamics of the parent-child relationship was surmounted by attachment theory's focus on the importance of measureable behaviors such as expression of negative emotions, violence, and abandonment.

Self-Efficacy Theory. This theory describes the factors that impact an individual's belief that he or she can carry out a required behavior to produce a desired outcome (Bandura, 1977). Self-efficacy, according to Bandura (1997), is the exercise of control over the events that affect one's life. In the context of maltreatment, self-efficacy relates to the ability of the caregiver to develop the necessary parenting skills and manage the resources required to adequately and

safely care for a child. Individual differences in motivation for change through actionable behaviors are a function of individual beliefs and cognitive processes. This theory is quite relevant to the development of maltreatment prevention programs that work directly with caregivers through parenting education and skill development.

Self-efficacy theory distinguishes between "efficacy" expectations and "outcome" expectations. Outcome expectations are the belief that a given behavior will produce a required outcome (*"learning parenting skills will make me a better parent"*). Efficacy expectations are an individuals' belief that he or she can actually initiate and carry out that required behavior (*"I can learn the parenting skills and then follow through with them with my own children"*). These individual beliefs play a major role in the level of personal effort that will be expended to attempt to resolve a given challenge.

Bandura (1994) outlines four sources of influence that lead to the development of a person's belief about their own efficacy and that can be directly applied to the development of interventions that target self-efficacy. Specifically, these points are pertinent in the development of programs and interventions that seek to improve a parent's belief that they can provide appropriate and nurturing care for their child. First, and most effective, is to develop efficacy through mastery experiences. A strong and resilient sense of success can be achieved through perseverant effort in overcoming challenging obstacles. This approach comes through gradually building up an individual's confidence that he or she can carry out tasks in which they previously would have failed. Once an individual has developed a sense that they have the ability to overcome adversity, especially during difficult times, their confidence and ability to execute tasks is stronger.

Second, self-efficacy can be strengthened through vicarious experiences provided through social modeling. Watching another person persist despite struggle in a similar circumstance instills the notion that they also have the ability to master comparable challenges. The success of this approach is strongly influenced by the perceived sense of success in the model and the similarities between the individual and model's circumstances. Models not only provide inspiration as a comparative social standard, they also teach necessary skills and strategies for managing challenges. In a similar vein, the third method of developing self-efficacy is social persuasion. Use of verbal praise and affirmation from an outside party that the individual has the abilities to master activities can promote stronger effort and extinguish self-doubt. Social persuasion is often more powerful working in the negative direction. Thus, external sources of persuasion must structure situations that bring success and avoid failure.

The fourth way to build self-efficacy is to reduce the natural stress reactions and negative emotional states that typically accompany adversity and vulnerability. People interpret the stress, fatigue, pain, and hopeless mood that accompany the demands of a difficult life as a sign of emotional weakness and physical incapacity. The key to promoting self-efficacy in this way is to change the way that an individual perceives and reacts to the emotional and physical reactions from self-doubt to one of excitement and energy.

Summary of Theory

Interventions to improve parenting skills, including home visiting, are making the transition to centrally placing theory and empiricism as primary drivers of program development. This is at least partially due to the evidence-based practice movement in medicine and more recently in mental health and social services (Thyer, 2008). There is now an evolving effort to develop a "science" for effective parenting interventions to meet the high demand and need for

these types of services. As the most common service provided, over half of families receive some type of parent training following a child welfare investigation (Hurlburt, Barth, Leslie, Landsverk, & McCrae, 2007). Over the past decade, there has been a significant increase in the supply of parent training interventions with empirical support. While the general approach to parent-focused prevention has been consistent, interventions have evolved over recent years from ideology-based to evidence-based (Institute of Medicine and National Resource Council, 2012). The background for this shift from doing what we think might work, to what we can support with theory and evidence is an important backdrop for this chapter.

Model developers are consistently exploring theory in empirical papers and one developer describes their model as "research-based and theory-driven" (Olds, 1999). Surely an understanding of theories of child development and caregiver behavior change in the context of human ecology is an important starting point for developing a maltreatment prevention strategy. For example, self-efficacy theory should inform interventions that teach parenting skills, particularly among caregivers who lack motivation to change. Simply providing at-risk parents with knowledge about developmental stages is not sufficient. Successful interventions also recognize that developing self-efficacy is critical given that many parents have not had the benefit of a stable family environment in their own pasts and often lack the social support and models for good parenting behavior. This theory was incorporated into the design of NFP by focusing on improving the confidence of mothers by setting small, achievable goals to promote the desired long-term changes (Olds, Kitzman, Cole, & Robinson, 1997).

The next section describes the state of the empirical literature advancing early childhood home visiting. There has been a great deal of opportunity to test the theories of change of home visiting for the prevention of child maltreatment. The section will first begin with an overview

of how home visiting is identified as an intervention and then describe the body of knowledge that supports claims of effectiveness.

Empirical Evidence of Home Visiting Effectiveness

Prenatal and intensive postnatal visitation for first-time mothers with services focusing on healthy child development is a potentially beneficial strategy to foster healthy attachment and reduce the risk for maltreatment (MacMillan, 2009). The CDC's Task Force on Community Preventive Services (Biluhka, et al., 2005, p. 11) defined "home visitation" as:

> "a program that includes visitation of parent(s) and child(ren) in their home by trained personnel who convey information about child health, development, and care; offer support; provide training; or deliver any combination of these services. Visits must occur during at least part of the child's first 2 years of life, but can begin during pregnancy and can continue after the child's second birthday. Programs may include (but are not limited to) one or more of the following components: training of parent(s) on prenatal and infant care; training on parenting to prevent child abuse and neglect; developmental interaction with infants and toddlers; family planning assistance; development of problem-solving and life skills; educational and work opportunities; and linkage with community services."

Home visitation is simply a method of service delivery, therefore there is much variation across early childhood home-visiting programs on theoretical approach, target family, services offered, home visitor role, and program model (Howard & Brooks-Gunn, 2009). Home visiting programs targeting new mothers often have specified outcomes other than maltreatment prevention but most center on maternal and child health. A systematic review of all home visitation programs reported a reduction of reported maltreatment by 39% with visitation by nurses or mental health workers yielding greater effects than by paraprofessionals (Bilukha, et al., 2005). The US Department of Health and Human Services recently commissioned the Home Visiting Evidence of Effectiveness evaluation to review the literature and examine the effectiveness and implementation guidelines of all existing home-visiting programs (Paulsell, Avellar, Martin, & Grosso, 2010). The models identified as having a sufficient evidence-base

for federal funding were Nurse-Family Partnership (NFP), Early Head Start, Family Check-Up, Healthy Families America, Healthy Start, Healthy Steps, Resource Mothers Program, Parents as Teachers, Home Instruction for Parents of Preschool Youngsters, and SafeCare. The list of evidence-based models will be regularly reviewed and updated as promising programs develop evaluation strategies and publish findings in peer-reviewed outlets.

Each model was examined across a set of criteria to determine if it has: been in existence for at least three years, an association with a national organization or institute of higher education, a minimum number of visits, a minimum education requirements for staff, a supervision requirement for home visitors, pre-service training requirements, fidelity standards local agencies must follow, a system for monitoring fidelity, and specified content and activities for home visits. While, all of these models have been in existence for at least three years, NFP was the only model to achieve all of the criteria set forth by the evaluators.

Nurse-Family Partnership. The most "mature" home visiting program in the United States is NFP. This section will describe this intervention as it was used to directly inform the Nurses for Newborns model that is the subject of this study and because it is in a position to likely be the most widely implemented home visiting model in the MIECHV expansion. One variation of the home visitation model, NFP's popularity and support has been bolstered by the rigorous and extensive longitudinal evaluation of outcomes (Scribano, 2010; MacMillan, 2009). NFP has been tested in three randomized control trials across a fairly diverse range of samples and demographic regions. The goals of NFP are quite ambitious, seeking to alter the "adverse maternal health-related behaviors during pregnancy, compromised care of the child, and stressful conditions in families' homes" and thereby prevent "the most pervasive and intractable problems faced by young children and parents" (Olds, et al., 1999). To help caregivers provide adequate

physical care, parents are taught to monitor signs of illness and ensure child safety in the household environment. To enhance emotional care for the child, nurses deliver information aimed to develop a positive parent-child interaction style and promote development through stimulating play.

The NFP model consists of home visits by registered nurses to first-time mothers beginning prenatally and continuing until the child's second birthday. The frequency of visits is set by the stages of pregnancy and childhood with some leeway to adapt to the family's needs. First time mothers are enrolled during the second trimester and visits occur weekly for one month. Visits are then scheduled for every other week until birth, then postnatal visits become weekly for six more weeks. Visits occur twice a month from two to 21 months, and then are tapered to once a month until the second birthday. Mothers in the first trial had an average of nine prenatal and 23 postnatal visits, with variation between mothers. Nurses carry a caseload of 20 to 25 families and visits last 75 to 90 minutes. The three major aims of the activities in the visits in the NFP model (Olds, et al,. 1999, p. 49) are to: 1) promote improvements in women's (and other family members') behavior thought to affect pregnancy outcomes, the health and development of the children, and parents' life course, 2) help women build supportive relationships with family members and friends and 3) link women and their family members with other needed health and human services.

The first study site for NFP was in Elmira, New York. The design consisted of a 4-arm randomized controlled trial. 400 pregnant women with no previous live birth and at least one risk factor (less than 19 years old, single, low socioeconomic status) were randomized to four conditions which consist of increasing layers of service beginning with developmental screenings only, then free transportation for prenatal and well-child care, home visits during

pregnancy, and finally home visits during pregnancy and until second birthday. Data was collected during interviews, observations in the home, and reviews of administrative health and social service records with follow-up now at the child's 15th birthday.

By the child's second birthday the results showed a reduction in official abuse and neglect reports and emergency department visits compared to control groups (Olds, 1986; Olds, et al., 1997; Kitzman, et al. 1997). These findings include a marginally-significant 80% reduction in maltreatment among single, low-income, teen mothers (4% vs. 19%, $p = .07$). The 15-year follow-up indicated that the difference in official reports increased between the group receiving prenatal and postnatal visits compared to the controls (.29 reports per participants vs. .54, $p<.001$).

The study was then replicated with a lower-income sample of African-American women receiving services through an existing health department in Memphis, Tennessee. Another later trial in Denver, Colorado consisted of a large sample of Hispanics and had a specific focus on examining outcomes for nurse-visited mothers compared to those served by paraprofessionals instead of nurses (Olds, 2008). In both Memphis and Denver, the rate of state-verified reports of child abuse and neglect were too low (3-4%) to adequately address the impact of the program on maltreatment prevention. The Memphis trial had a 23% reduction in health-care encounters and 79% fewer hospitalizations for injuries compared to the control group. The most recent review of this study with children now at age 12 has demonstrated less use of cigarettes, alcohol, and marijuana and less prevalence of internalizing disorders compared to the control group (Kitzman, et al., 2010). Additionally, nurse-visited children scored higher on reading and math tests than the control group. An indicator of decreased maltreatment, children in the Denver trial at age 4

had improved developmental outcomes including advanced language, executive functioning, and behavioral adaptation (Olds, et al., 2004).

The benefits of NFP are also found in outcomes for the mothers, which provide many direct benefits to the child. As this program targets low-income and often single mothers, this intervention has potentially profound impacts on spending for programs designed to assist these families. Olds and colleagues (2010) found less impairment due to drug and alcohol use, longer partner relationships, and greater sense of mastery among the same mothers 12 years after the intervention. These changes translate to less government spending on food stamps, Medicaid, and Aid to Families with Dependent Children. These long-term savings are greater than the invested cost of the program per family. Two economic analyses (Aos, et al., 2004 and Karoly, et al, 2005) estimated an approximate savings of $17,000 per family across the three trials of NFP (Olds, 2010). Not only are these programs effective in reducing harmful outcomes for both the child and the parent, an important aspect of developing sound health policy is promoting programs that are proven to be cost effective.

The research on NFP and subsequent dissemination of the model has had a major impact on the field of maltreatment prevention and on policies that expand home visiting services. However, there are many limitations of the NFP model and gaps in the knowledge base when focusing solely on NFP studies. Many of these gaps and limitations are addressed by the NFN model and the disseminated research it provides. First, the focus of NFP is primarily on preparing first-time mothers with specific risk factors (low-income, unmarried) with the transition to caregiving. While this is certainly an important subpopulation on which to concentrate resources, it is quite narrowly focused and excludes many families in need. For example, a mother whose second child was born with significant medical complications could

still greatly benefit from nurse home-visiting services. However, the family would not be eligible for NFP. This represents a significant missed opportunity. Further, if one *could* provide NFP services to this mother, there is no indication this model would be the appropriate service approach. There is currently a need for services that can provide support to the families that fall outside the NFP criteria. NFN provides the opportunity to expand what we know to be possible in reaching and helping the diverse population of families with new babies in their home. The following table contrasts the NFN and NFP programs across key program components.

Table 1
Comparing and Contrasting NFN and NFP Home-Visiting Programs

	Nurses for Newborns (NFN)	Nurse-Family Partnership (NFP)
Program history	Founded in 1991 with mission to prevent infant mortality and child abuse and neglect	David Olds began randomized trials in 1977 and replication began in 1996
Geographic area served	25 counties in Missouri and 7 in Tennessee	Provided nationally in 32 states and 396 counties RCTs in New York, Tennessee, and Colorado and trials underway in Netherlands, UK, and Canada
Onset of Services	Referrals can be made prenatally and up to 18 months after the child is born	Referrals must be made prenatally by 28th week of pregnancy
Inclusion/ exclusion criteria	Any expectant or new mother that is referred for services is eligible	Only first time, low-income mothers enrolled prenatally
Duration of services	The minimum service level varies by program (6-10 visits) and case closure is determined by the nurse based on each family circumstance and needs	NFP nurses plan 64 total home visits with weekly home visits for the first month after enrollment and then every other week until the baby is born. Visits are weekly for the first six weeks after the baby is born, and then every other week until the baby is 20 months. The last four visits are monthly until the child is 2 years old.

Visitor information	Registered Nurses (RNs) with minimum of 3 years of experience	Registered professional nurses with minimum of BS in nursing
Population demographics	93% Medicaid eligible, 38% medically fragile babies, 34% medically challenged mom, 34% less than high school education	Nationally, median age 19, 88% unmarried, 51% completed high school, median income $13,500, Missouri NFP reported 91% of babies born full-term at healthy weight
Program Cost	$211.52 program services cost per visit	$4,500 per family per year (range $2,914-$6,463)

Gaps in the Literature

There is clearly more research needed to delineate the key components of an effective wide-reaching program that can be delivered with high levels of fidelity at a broad scale. Daro (2009) suggests that attributes of a successful home-visiting program include a well-articulated and specific theory of change, evidence of better outcomes as a consequence of program participation compared to those not receiving services, evidence of impacts on specific domains for specific participants, evidence of impact on the provider due to implementation challenges, and an established method of program replication.

Unfortunately, few programs move beyond their ability to articulate an initial model for change and provide initial evidence of better outcomes. There is a consistent theme of lack of positive findings in reducing child maltreatment in programs other than NFP. While there is some indication that positive results can be found in caregiver self-report or observations of the home environment, there is little evidence of improvement in actual rates of maltreatment. There is also a lack of research examining the question, "what works best for whom?" The dose and subpopulation analysis will provide critical information that has been lacking in research looking solely at average effects across diverse populations. As replication and dissemination continues,

areas will need further knowledge about how to implement a home-visiting program to best serve the needs of their target population. The more specific results that research can provide to inform best practice, the more effective and efficient tailored programs can be.

The current research study builds on the existing theoretical and empirical foundation that has been provided by thirty years of efforts in the field of maltreatment prevention through home-visiting services. The NFN model provides an opportunity to examine the effectiveness of a nurse home-visiting model building on the work of the NFP model. While there are similarities in the intervention model (nurse-delivered home visiting services) and in the outcomes examined (maltreatment prevention), the current study differs from the design of NFP RCTs in two ways. First, this study examines the effectiveness outside the context of a controlled trial to truly test the generalizability of this type of intervention in a community-based implementation. Second, the NFN program serves a wider range of higher-risk families than has been studied with the NFP model. These results provide much needed information about the potential of home-visiting services for very high-risk families. While the NFP studies have provided evidence regarding the efficacy of home visiting under ideal circumstances, this type of research is needed to understand what types of outcomes can be expected when services are delivered by agencies in the community with a population of families most in need of these services.

Summary of Research Aims Response to Known Gaps

Aim 1 directly addresses gaps in knowledge about nurse home visitation programs outside the NFP Olds model. There is little research on programs that deploy nurses for home visits yet maintain a flexible program curriculum and accept non-first time mothers for services. This aim seeks to determine what the extent to which this program model delivers different

levels of engagement and outcomes. Child physical health and maltreatment are the main outcomes of interest from the perspective of the NFN program. The focus of the present dissertation is on maltreatment prevention as measured by alleged reports as it was not possible to obtain a source of health data external to the program files.

Aim 2 examines major groups of interest that are relevant to the ideas of expanding home visitation as a national model for improving child health. The subgroups of interest were chosen to be relevant to policy and practice. For example, it is important to determine whether or not families who differ by geography and race have different outcomes when served by the same program. This relates to the cultural acceptability of a program and the need to potentially tailor services for families depending on where they live, the age of the mother, or their racial identity. There is also interest to understand whether targeted prevention programs (such as the Olds model) that require mothers to be primiparous and enroll prenatally, confer any measureable benefit by not accepting multiparous mothers and those who enroll in the postpartum period. Further there is a pressing need to understand how specific maternal and infant risk factors relate to differential outcomes. This analysis focuses on medically-fragile newborns and mothers with mental health and substance abuse concerns. Programs and policymakers need to know whether the same outcomes and level of engagement can be expected for different groups of families in order inform the planning and delivery of services to reach all families.

The purpose of Aim 3 is to assess questions of causation and program effectiveness. Does NFN cause a decreased risk for child abuse and neglect? While Aims 1 and 2 examine pre and post intervention outcomes, this type of analysis is strictly correlational and cannot test any causal claims. Like many programs developed in the community, the NFN program has been operating for a significant amount of time, taking any family that meets broad criteria and with

funding and saturation issues that preclude randomizing to services or no services. The final aim of this dissertation is to apply an advanced statistical technique, propensity score matching, to explore the ability to detect a program effect size using program drop-outs as controls.

Thus, this dissertation advances the field by thoroughly describing an alternate model of nurse home visitation for high risk families, exploring methodological issues in the measure of program dose as well as how to assess effect in the absence of ability to conduct a RCT, and finally in exploring how outcomes are associated with broader population groups that are relevant to thinking about nurse home visitation expansion.

Chapter 3: Research Design and Methodology

The following chapter provides a detailed description of the research and analytic strategy used to understand the role of nurse home visiting services provided by Nurses for Newborns (NFN) in the promotion of healthy infant development and the prevention of maltreatment. The research plan is intended to be consistent with other research on home visiting programs to allow for the discussion of results and implications in the broader policy context of home visiting service expansion. However, the research plan was also developed in close discussion with agency leaders to ensure that results were relevant to specific practice questions. The overall approach for the study will be discussed first, and then the methodology for each aim and associated research question will be described in detail.

Methods Overview

This study was a longitudinal investigation of the prevention of child maltreatment among families receiving services from a well-established nurse home-visiting program. Although the services and outcomes have already occurred, exact dates collected from electronic services data allowed for prospective analyses. All NFN nurses carry laptops to home visits used to document case notes at each visit and to collect information in areas such as infant health, child development, healthcare use, and family risk factors. Each laptop has an automatic link to the central database at the agency where data are stored. This archived data allowed for a unique opportunity to conduct longitudinal research with a relatively large sample in a short amount of time. The other source of data for this study came from linked administrative data from the state child welfare child abuse and neglect (CA/N) reporting data.

NFN employs Registered Nurses with at least three years of experience in a special care nursery or neonatal intensive care unit, or five years of experience in community health or

maternal health. Nurse training includes a classroom component covering the clinical guidelines and a mentoring process for the first three months prior to independent home visits. Caseloads range from 25 to 40 families. Based on needs assessments and screening outcomes, nurses present educational topics designed to address identified issues and improve parenting skills and help parents establish access to external resources to promote the health and safety of both the mother and child. Although the program is designed to be flexible, fidelity to clinical guidelines is accomplished through weekly clinical case reviews, monthly case management meetings, peer and supervisory reviews, and client satisfaction surveys.

Program Participants/Study Sample

While NFN operates in two states, Tennessee and Missouri, the present study sample is drawn from families receiving NFN services from the St. Louis-based Missouri program. NFN accepts referrals anytime from prenatal visits up to 18 months after birth. Almost all cases served are closed by age two. Most of the NFN referrals originate from social workers and other hospital staff based on a perceived risk of maltreatment (Jonson-Reid & Stahlschmidt, 2010). For example, pregnant mothers or those with newborns that have been referred to the child welfare system that do not yet meet state definitions for investigation or assessment of maltreatment are often referred to NFN for community-based preventive services. Over 85% of all the families served by the program qualify for Medicaid based on income with even greater coverage given the expanded eligibility for infants up to 300% FPL (Jonson-Reid & Stahlschmidt, 2010). Due to the demographics of low-income families within this region, 93% of the families were either African-American or non-Hispanic White race/ethnic categories. While NFN has been in operation for over 20 years, their target population differs substantively

from other nurse home visiting programs, but has not received the same amount of research attention. The first part of Aim 1 is to describe their service population.

Inclusion/exclusion criteria. Recently, NFN updated their practice to reflect changes in the field of evidence-based home visiting to include implementation of validated screening tools and revised clinical guidelines. In order to have consistent programming and to take advantage of the inclusion of the new tools, it was determined that the retrospective case review entry point should begin in 2009. Cases were further limited to those closed by July of 2011 to allow for an adequate sample size and a sufficient follow-up period to examine the child maltreatment report outcome.

Based on this study time period, the original sampling frame included 7,154 families. Because baseline assessment information was critical to the study, cases that were referred but never received an initial visit were excluded. To remain consistent with the federal home visiting expansion focus on low-income families, only families eligible or receiving Medicaid, WIC, or TANF were included. Some families had more than one child receiving services during the service period, so the child with the most documented visits within that family was selected as the target child to track outcomes. This eliminated potential bias due to clustering at the family level. These further sample restrictions yielded 3,620 families providing adequate sample size to conduct multivariate analysis, propensity score matching, and subgroup analyses with sufficient elapsed time to examine prevention of subsequent maltreatment reports. A power analysis was conducted based on an alpha = .05, a power = .80, and assuming a small effect size (d = .20) based on a meta-analysis (Geeraert, et al., 2004) and indicated that a minimum of 393 subjects were needed to detect an effect for official reports.

Measurement

All independent variables including child and family demographic, risk factors, and service utilization came from agency case records. The primary outcome of interest, prevention of child maltreatment, was determined by linking NFN data to the state-level child maltreatment reporting data to see if families were reported for alleged abuse or neglect. Linking to the state child welfare was completed using a common child case identifier or Social Security number for the majority of cases. If this was not available then the child was matched using the first four letters of the first name, the first four letters of the last name, and the child's date of birth.

Not only are the variables consistent with constructs explored in prior home visiting research, they were also cross-referenced, whenever possible, with federal benchmarks. This policy framework provides a means of comparing findings from the present study to what is being expected for programs receiving federal MIECHV funds through state block grants. The next section describes in further detail the independent and dependent variables for this study.

Independent variables: Family demographic information. NFN agency data included a summary case file and detailed information for each visit. The following variables were included in the existing NFN administrative data collected by nurses at referral and the first assessment home visit. Information about the child and parent was used as covariates, predictors, or grouping variables depending on the research question and analysis.

Maternal and child demographic characteristics. About 98% of caregivers and 92% of infants in the sample were documented as either "Caucasian" or "Black". Child's race initially included the categories of "Pacific Islander", "American Indian", and "Asian" which were less than one percent of the sample and about 7% percent were listed as "Biracial". Dealing with smaller race subgroups presents a significant challenge. One option would be to eliminate all races other than White or African-American from the sample. Since they represent such a small

portion of the sample, there would likely be issues with small cell sizes in bivariate analyses and separation in regression analyses. However, deleting these families from the study due to a single variable is also problematic given the focus of the study on representing an existing program in practice. Based on preliminary analyses examining differences in demographic and outcome variables across race groups it was determined that for multivariate analyses in Aim 1 and Aim 3 race was dichotomized as African-American/not African-American. In Aim 2, there is a specific focus on subgroup analysis for race groups and thus the original race coding was maintained except that "Pacific Islander", "American Indian", and "Asian" were collapsed into one group. Additionally, about 4% of children were coded as "Hispanic" and ethnicity coded as Hispanic/non-Hispanic.

Child gender was retained as either "male" or "female". The birthweight of the child was coded as "healthy weight", "low birthweight (LBW)" (<2500 grams), and "very low birthweight (VLBW)" (<1500 grams). Children born less than 37 weeks gestation were coded as "preterm birth" or "term birth" for those at 37 weeks or more gestation. These medical definitions for thresholds of birthweight and gestation are widely used in the medical literature and are consistent with the Healthy People 2020 Maternal, Infant, and Child Health indicators (DHHS, 2010) and the Missouri State needs assessment operationalization of identified MIECHV benchmarks (Missouri Department of Health and Senior Services, nd).

A distinction should be noted in the term "maternal" versus "caregiver" in this study. The biological mother of the child is primary caregiver in 97% of the cases. Thus, in most cases, these refer to the same individual. However, in some cases a maternal risk factor refers to the biological mother (alcohol use during pregnancy) while a caregiver demographic characteristic (education) refers to a different individual who is actually the primary caregiver. In these cases,

the biological mother is not the one caring for the child, thus this identified caregiver is the one directly receiving the home visiting services. Instead of deleting these cases, a dichotomous variable was created to capture the caregiver's relationships to infant as biological mother/not biological mother. For this reason, the term caregiver is used to describe the individual caring for the infant in the postpartum period (not always the biological mother) while maternal is used to describe the individual carrying the child prenatally (always the biological mother).

Maternal pregnancy history was captured using nurse documentation of the commonly-used documentation of gravida/para/living. This notation provides the obstetric history of how many times the mother has been pregnant (gravida), the number of births beyond 20 weeks (para), and the number of living children (living). Additionally, a risk factor was created for history of prior fetal death to reflect nurse documentation of prior miscarriages, spontaneous abortions, or stillbirths that may not have been included in the obstetric history count. This documentation was used to create variables for the subgroup aim examining outcomes for primiparous (first-time) and multiparous mothers as well as in the risk factor coding for families with multiple other children in the home.

Socioeconomic characteristics of the caregiver included marital status which initially included seven categories of married, single, widowed, separated, divorced, consensual union, or other. These were re-coded as partner (married or consensual union)/no partner (other groups). Caregiver education ranged from no high school, GED, high school degree, some college, Associate's degree, Bachelor's degree, and Graduate degree. This was recoded into four categories for bivariate analyses as no high school, high school or GED, some college, and college degree. The no high school group was the largest in the sample and to reflect the high level of risk associated with a lack of high school degree, this variable was recoded as a

dichotomous risk factor as at least high school/no high school education. Caregiver employment status included categories of unemployed, disabled, homemaker, and student, part-time, full-time, and other. This variable was recoded as a dichotomous risk factor as unemployed/not unemployed. While this recoding has the limitation of placing homemakers, students, and employed individuals in the same group, this grouping reflects families that likely experience a lower level of risk and psychosocial stress for those who are in the employed or are not in the workforce due to other responsibilities (homemakers/students) compared to those who are unemployed but are able to participate in the workforce.

Geographic location was determined using the zip code of the family. Several zip-level independent variables were created for this study. Three categories were created to determine urbanicity of the family's home location categorized as rural, urban, or suburban. Zip codes for families were merged with the publicly-available Rural-Urban Commuting Area (RUCA) Codes. This is a classification scheme developed by the Census Bureau to classify year 2000 Census tract-level data and 2004 zip codes regarding their urban and rural status. Additionally, zip-code level demographic and social indicators from the 2000 US Census were linked with the existing dataset.

An exploratory correlational analysis was completed using two outcomes of interest, number of visits and child CA/N report, and 14 zip-level Census indicators (population density, percent married, percent females with no high school, percent female not in the labor force, percent female unemployment, percent with more than an hour commute, per capita income, percent families renting, median household income, median family income, percent individuals using SSI, percent individuals using TANF, and child poverty rate). To avoid issues of multicollinearity by including these highly correlated variables in a single model, two variables

were selected for testing in final models, median family income, and child poverty, as they were highly correlated with the key outcomes as well as other zip-level indicators. Given the high correlation between the two variables, they were tested separately for significance in main models. These variables were not significant in any final models and so were both kept in final models as control variables.

Independent variables: caregiver and child risk factors. NFN collects information on specific risk factors at intake. Risk factors are coded as yes/no dichotomous variables and included the following 13 pre-defined areas in the agency dataset: alcohol abuse during pregnancy, drug abuse during pregnancy, baby medical problems, maternal medical problems, "challenged" mother, caregiver history of abuse, problems during pregnancy, inadequate prenatal care, late prenatal care, less than 18 months between pregnancies, psychological/social risk, caregiver smoking, and household smoking. In addition to these dichotomous risk variables, the nurse assessment notes identify additional details about these risk factors. For example, if a nurse checks mental health problem, the notes might list a specific diagnosis. Notes from the child and maternal risk factor categories were used to measure additional risk factors identified in the literature or clarify existing ones across areas of psychosocial risk, violence exposure, behavioral health, and maternal health.

A final set of 31 variables measuring specific risk factors were summed to create a cumulative risk score that was used as a dimensional measure of overall level of risk. Cumulative risk modeling is a measurement approach applied to complex systems that simultaneously accounts for multiple risks in predicting an outcome while reflecting the natural covariation of a set of measured risk factors (Evans, 2003). Grounded in the work of Sameroff (Sameroff, Seifer, Barocas, Zax, & Greenspan, 1987; Sameroff., 1998) and Rutter (1979),

proponents of cumulative risk modeling suggest that no single risk factor can reliably predict most outcomes of interest. Further, use of a cumulative risk score negates the potential for multicollinearity when modeling unique risk factors that are highly associated (MacKenzie, et al., 2011).

It must be highlighted that these risk measures are almost entirely drawn from self-report from the caregiver and the nurse. The nurse provides a clinical interview and documents the risk factors that he or she determines are most relevant based on the information caregiver provides. Additional information is often gained through the referral process from other providers that may also be contained in this information.

Child-level risk factors. Nurse notes were reviewed to determine what types of issues were commonly commented on by the nurse as an indication of elevated risk. Other risk variables were created based on known literature on factors that increase risk for child maltreatment. Based on the linked CA/N administrative data, a risk variable was created to measure children who were reported victim of a maltreatment referral prior to NFN services. A risk variable was created to measure children who were identified as having a documented stay in a special care unit or neonatal intensive care unit (NICU) and for those that were developmentally small during the perinatal period due to low birthweight, preterm birth, intrauterine growth restriction, or slow postnatal weight gain. A unique risk variable was created to measure children with health issues at birth documenting drug-exposure or withdrawal, jaundice or hyperbilirubinemia, significant heart or lung issues (respiratory distress, pneumonia, bradycardia, etc.), and major congenital disabilities or abnormalities presenting at birth (teratomas, genetic disorders, deformities, seizures, etc.). For the Aim 2 subgroup analysis, an

indicator for "medically-fragile" infants was created for those with any of the risk factors for special care/NICU, small baby, drug exposed, cardiorespiratory issues, or congenital disability.

Caregiver psychosocial risk factors. From the nurse notes, a risk variable was created to measure mothers who were documented as being homeless or having an inadequate or unstable living situation. A risk variable measured whether the paternity of the child was not clear or the father of the baby was incarcerated, had died, was deployed in the military, was unsupportive of the mother, or was for another reason not involved. A risk variable was developed to identify cases where the mother had prior involvement with the Division of Family Services (DFS) through loss of custody of a child, having parental rights terminated, or a self-reported prior hotline call. Risk variables identifying teenage mothers, caregivers with no high school education, those with multiple current children in the home, and unemployed caregivers were also created based on nurse documentation. A separate section for psychosocial risk factors pertaining to mental health or substance abuse is discussed below.

Caregiver violence exposure risk factors. A set of risk variables measuring the caregiver's history of violence exposure was developed using the nurse comments. This includes measuring current partner or domestic violence from report of an abusive relationship, history of sexual abuse or rape, and reported history of child maltreatment. It should again be noted that the presence of these risk factors is solely based on self-report of the mother and then on documentation of the nurse. This type of measurement likely results in significant under-reporting of violence exposure in this sample.

Caregiver behavioral health risk factors. These risk variables focus on measuring the presence of issues of caregiver substance abuse and mental health. Given the focus of the agency on high-risk mothers, there is a wide variety of overlapping issues that cover maternal

behavioral health risk. A risk variable was created for mothers with self-reported alcohol use during pregnancy and for drug use during pregnancy. A mental health risk indicator included mothers with a self-reported diagnosis of a mood or behavioral disorder such as depression, anxiety disorder, or bipolar disorder or those with reported suicide attempts. Mothers with a documented learning disorder, cognitive or developmental disability, or who receive special education services in school were categorized with cognitive/learning disabilities. A risk variable was created for mothers with neurologic impairment or injuries including traumatic brain injury, seizure disorder, and epilepsy. For Aim 2 examining subgroup outcomes, an indicator was developed for mothers with a mental health disorder alone, a substance abuse disorder alone, and those with co-occurring mental health and substance abuse issues.

Maternal health risk factors. Several risk variables were created for maternal health factors related to prenatal health risk. The nurses documented whether or not the pregnancy was intended. A risk variable was created based on nurse notes to identify mothers with prior low birthweight or preterm births, prior fetal deaths, less than 18 months between pregnancies, and inadequate or late prenatal care. A risk variable indicated mothers with prior or current pregnancy complications including preeclampsia, gestational diabetes, and pregnancy-induced hypertension. A risk variable was also created indicating mothers with other chronic physical health problems including chronic cardiorespiratory issues such as asthma and hypertension, sexually transmitted infections, and other major physical disorders.

Validated screening tools. One of the updates to the clinical process utilized by NFN was to add validated screening tools that could be given during home visits to assess for caregiver stress, maternal depression, and child developmental needs. These measures are used as either explanatory independent variables or outcome dependent variables depending on the

research aim and specific question. Maternal depression and stress as well as child developmental delays are known be strong risk factors for child maltreatment and are therefore important predictors for this outcome (Kotch, Browne, Dufort, Winsor, & Catellier, 1999; Taylor, Guterman, Lee, & Rathouz, 2009).

Edinburgh Postnatal Depression Scale (EPDS; Cox, Holden, & Sagovsky, 1987). Nurses used the EPDS to screen for postpartum depression and make referrals to mental health services. The EPDS is a brief screening tool specifically created for postnatal depression that has been validated for use with pregnant mothers (Murray & Cox, 1990). The scale uses 10 self-rated statements related to symptoms of depression with a four point (0-3) Likert-type response for each item. Items include prompts such as "I have felt sad or miserable", "I have looked forward with enjoyment to things", and "I have blamed myself unnecessarily when things went wrong". Scores range from 0-30 with a higher score indicating a greater severity of symptoms. A score of >10 indicates possible depression and a score of >13 indicating likelihood of depression. The scale has been found to have a third-grade reading level (Logsdon and Hutti, 2006) and an internal consistency ranging from .73 to .87 across seven studies (Boyd, et al., 2005).

Factor analyses have found that the EPDS may actually measure multidimensional constructs including anxiety, anhedonia, and suicidality as opposed to a unidimensional measure of depression (Phillips, et al., 2005; King, 2012). A review of 18 validation studies found relatively wide confidence intervals for sensitivity (65-100%) and specificity (49-100%; Eberhard-Gran, et al., 2001). Other studies have found lower reliability during the first two weeks after pregnancy due to unstable scores (Sheeder, Kabir & Stafford, 2009). There is also some question regarding the most appropriate cutoff score for different populations (Freed,

Chan, Boger, & Thompson, 2012). One study found that in a low-income urban population, a cutoff score of 7 was more appropriate (Chaudron, et al., 2007).

Everyday Stress Index (ESI; Hall, 1985). The ESI was originally designed for use with low income mothers with young children and is used by the agency to screen for chronic daily stress. The instrument uses 20 items across domains of financial concerns, role overload, employment problems, parenting worries, and interpersonal conflict using a four point (0-3) Likert-type scale to assess the extent to which caregivers are worried, upset, or bothered on a daily basis. The summary score ranges from 0-60 with higher scores indicating higher stress with no pre-defined threshold to identify a caregiver who is in a problem range. Internal consistency ranged from .80-.85 in previous studies (Hall, Gurley, Sachs, & Kryscio, 1991; Hall, Williams, & Greenberg, 1985; Hatcher, Rayens, Peden, & Hall, 2008) and the scale has been used in studies predicting risk for child maltreatment (Kotch, et al., 1995). For this study and most studies, the ESI is used as a continuous variable in analysis. In the subgroup analysis for this study, a "high stress" group threshold was created using the sample ESI mean of 9.15.

Measures of service utilization. Aim 1 of this study was centrally concerned with describing the service utilization of families receiving NFN home visiting services. This can be conceptualized as an outcome (e.g., engaged or not, completed program or not) or as a mediating or moderating variable (e.g., more services lead to improved child outcomes) depending on the research question. Analyses of time from referral to initiation of visits, initial engagement, retention in services and type of termination were examined. This section will describe how these different patterns of service use are defined.

Service referral and initial engagement. Service utilization constructs concerned with home visiting engagement are not consistently defined in the literature, thus some aspects of this

analysis can be considered exploratory. The first stage of service utilization measured is the time in days between when the referral is received by the agency and the first home visit. This time period was recoded as 1 day, 2-6 days, 7-30 days, and 30 or more days. This measure or "response time" gives an indication of the service system's ability to respond to needs of the family in a timely manner as well as perhaps the receptivity of the caregiver to the service.

Given the flexible period of enrollment and initiation of services, it is also critical to determine at what point in the prenatal or postpartum period services begin. A categorical variable was created based on the time between the child's date of birth and referral to include prenatal referral > 90 days before birth, prenatal referral < 90 days before birth, postnatal referral <30 days after birth, postnatal referral 30-90 days after birth, and postnatal referral > 90 days after birth. These categories were chosen to not only capture if the referral was prenatal or postpartum but also to describe the proximity of the referral to the event of the child's birth.

Families are referred from a variety of sources. The specific referral source for each family is captured in the agency documentation and was initially recoded into thirteen groups including self-referral, NFN referral, hospital, behavioral health, DFS/Social Services, pediatric clinic/doctor's office, high school, Head Start, public health office, other home visiting program, other social service, public insurance or managed care, and faith-based organization. This was recoded to include three groups: self-referral, healthcare providers, or other social service/education providers.

Families referred for services receive an initial home visit from the nurse for assessment and initial service planning. Among families in this sample, there are no cases where one visit would be deemed sufficient for service completion. At this point in a voluntary service, the caregiver can choose to continue services by allowing the nurse back into the home for a second

visit or they can passively refuse services by simply missing future appointments or not returning calls, or actively refuse services. In this study, an initially "engaged" caregiver is one who receives at least two home visits. This engagement indicator will be discussed further in Aim 3 as it is used to create the comparison group to assess program effectiveness.

Service retention. Hypothetically engagement with the family must be maintained over a sufficient period of time for the benefits of the program to accrue. Adequate length of services is termed "retention" in this study. There are several ways that this was measured. First, the raw number of visits as documented in the case notes was determined. Visits were categorized as prenatal, postpartum, or total home visits. Second, the total number of hours of service contact, also documented for each visit in the case notes was calculated. Last, the number of weeks from the first visit to the last visit was determined using dates of visits.

The last stage of service utilization is service termination. Categories describing the reason for termination based on existing NFN codes included refused services, family moved or nurse cannot locate, no need or duplication of services, completion of two years of visits, or other reason. Over half of the families were coded as "other" for their termination reason. Based on examination of this coding structure and consultation with the agency, it became clear that the termination notes were not coded consistently.

Delivery of information from nurse to caregiver. Another measure of service consumption is the delivery of pre-defined "teaching skills" tracked by the nurse. These skills are the count of informational/psychoeducational topics delivered from the nurse to the caregiver during the home visits. The agency has identified 57 total skills across areas including infant health and safety, injury prevention, child development, maternal health, and social support. The number of skills that a caregiver receives is determined by the relevance of the skill to their

needs and also the ability of the nurse to cover additional skills based on the parent's competence. The teaching skills serve as a proxy measure of the dose of specific service components that the caregiver receives. Teaching skills are used as an outcome of service provision but also a potential predictor of child development and maltreatment outcomes.

Dependent variable: child maltreatment reports. Child abuse and neglect reports were identified using linked state Children's Division (CD) data. Administrative data allows for the precise dating of an allegation of abuse or neglect, avoiding the difficulties in retrospective recall of maltreatment or services encountered in the past (Widom, Raphael & Dumont, 2004). Data from NFN were linked with CD data based on personal identifiers of the child (DCN, SSN, name, date of birth). In addition to the date of the report, the file included categorized reports regarding multiple response system track assignment (assessment or investigation), categories of alleged maltreatment type (medical abuse, neglect, physical abuse, sexual abuse), categories of substantiated maltreatment type (physical abuse, neglect, emotional abuse, medical neglect, educational neglect, sexual abuse, none – home schooling, and none), the severity of abuse (fatal, mild, moderate, permanent damage, serious/severe, unknown), 23 options of relationship of abuser to the child, and 17 options for reporter's job, and the conclusion code documenting whether or not the case was substantiated, unsubstantiated, or family assessment services were offered.

Given the confluence of empirical research suggesting that the "substantiation" label is not a useful predictor of risk for later harm or recidivism, primary comparisons were made between those families with a later report and those without a report regardless of case disposition (IOM and NRC, 2012; Kohl, Jonson-Reid, & Drake, 2009). However, the substantiation rate is reported for comparison with other studies and as a marker of burden on

child welfare and legal systems. It is also possible that a CA/N report occurs for reasons outside of the preventive scope of the NFN services. For example, sexual abuse by another family member with access to the child that is unknown to the caregiver may not be preventable with home visiting services. Thus, a more stringent "qualified" CA/N referral outcome was created limited to reports that have the parent as the perpetrator, sexual abuse is not the type, and the conclusion is not unsubstantiated. In some cases, CA/N reports occurred prior to the first home visit or between the first visit and service termination. Thus, only reports that occurred after services were terminated were classified as an adverse outcome for maltreatment prevention.

Agency data were matched with state child welfare administrative records in December 2012. Families began terminating from services in this study in January 2009 and the last family terminated services in March 2012. Therefore, the time to follow-up ranged from 10-47 months. While it is impossible to ensure that secondary administrative records are complete accounts of service history, the CA/N administrative data set has been used previously to follow families (Drake, Jonson-Reid, Way, & Chung, 2003) and prior researchers were able to confirm fewer than five percent of the identified families to have moved out of the region over a 7 year period. CA/N data are also statewide meaning that only families that moved out of state would be missed.

CPS reports only reflect those cases of maltreatment that are reported to authorities, greatly underestimating the true number of unreported cases. Further, surveillance bias in home visiting services is often listed as a problematic limitation of relying solely on official reports (Howard & Brooks-Gunn, 2009; Gomby, 2005). This limitation is minimized in the present study by having exact dates of reporting allowing for separate consideration of reports that occur

before, during or after services. Those occurring after termination can no longer be subject to this bias from this agency.

Dependent variable: child developmental outcomes. Outcomes from an evidence-based developmental screening tool, the Ages and Stages Questionnaire (ASQ; Squires, Bricker & Potter, 1997), were examined as an indicator of child health and development. The ASQ is used to identify infants and young children who need further evaluation or may require services for developmental delays or disorders (Squires, et al., 1999). One study found over 30% of zero to three year olds under investigation for maltreatment had ASQ scores above threshold for early intervention services (Casanueva, Cross, & Ringeisen, 2008). Another found 22% of substantiated cases scored in the problem-range of at least one developmental area (McCrae, Cahalane, Fusco, 2011). Another found 50% of children recently admitted to foster care had problem-range scores on the ASQ (Jee at al., 2010).

The ASQ assessment begins at four months of age. The assessment is completed by the parents and takes 10-15 minutes to complete covering five developmental areas: communication, gross motor, fine motor, problem solving, and personal-social. Parents rate whether or not their child performs a specific skill (10), sometimes performs the skill (5), or is not yet performing the skill (0). After scoring, children less than two standard deviations below the mean are considered in need of evaluation. In this analysis, children were identified as having problem-range scores or not at each visit the ASQ was provided. Change over time for this scale was assessed for families that completed the assessment at subsequent visits.

One limitation of this measure is the use of parent report of a child's developmental status. In a sample of children involved with child welfare, the NSCAW study found that caregivers had good specificity (85% accurately identified child did not have a delay) but poor

sensitivity (35% accurately identified child who did have a delay) in reporting their child's needs for developmental services (Berkoff, Leslie, & Stahmer, 2006). Families experiencing high stress when caring for an infant might have difficult accurately reporting this type of information (McCrae, Cahalane, & Fusco, 2011). However, data is collected by NFN every three months improving the ability to obtain reliable responses of concerns relating to child development.

Data Management

All data for this analysis is secondary administrative data originating from the NFN agency electronic data system and from state services data systems. Nurses enter the data directly into a laptop with programmed data input fields which are then maintained on a central agency server. Once the sampling frame and required variables were determined, all datasets were compiled by the NFN Director of Research, de-identified and assigned a unique case identifier, and transferred to the researcher. The Director of Nursing and Chief Executive Officer were available regularly for consultation regarding coding of variables and data collection process. The Director of Research made the data request for the state administrative data and maintained the identified datasets on the agency server. Similarly, the linked administrative data set was merged with the agency records and transferred to the researcher for analysis. All procedures were approved by the Washington University Human Protection Research Office.

Missing data strategy. Given the use of agency data and secondary administrative data, the current study has what could be considered a relatively low level of missing data for most measures. Since data is collected systematically during intake and assessment on the mother, only 1-3% of data is missing on maternal demographic characteristics and service dates. Based on the coding scheme of the risk factors, there is no data missing for these variables. The bulk of

the missing data in this study is found in child characteristics and the screening tools due largely to attrition. First, child demographic information was not collected on about 7% of cases because the caregiver dropped out of services before the child was born or before their information could be attained. For these families, there was no child date of birth or other identifying information and therefore the CA/N administrative data could not be linked. These cases were simply dropped from outcome analysis.

Clinical screening measures (EPDS = 56%, ESI = 46%, ASQ = 74%) were not collected largely as a result of attrition and present an analytic challenge in this study due to a significant threat to internal validity (Kristman, Manno, & Cote, 2005). Essentially, in most cases families dropped out of services before the nurse could deliver the screening tool. Previously, researchers would often delete cases with missing values or use other crude *ad hoc* imputation strategies that are more likely to yield biased results, result in a loss of statistical power, and have less theoretical support (Enders, 2010; Groenwald, Donders, Roes, Harrell, & Moons, 2012). Given that the probability of missing data in this study is related to other observed covariates, they are assumed to be missing at random (MAR) and more advanced techniques can be employed (Little & Rubin, 1987). In addition to understanding the mechanism, the pattern of missing data is also important in determining the approach to dealing with missing data. Missing longitudinal data with a monotone, non-arbitrary pattern allows for the use of a more flexible and theoretically-valid method that applies a sequential approach to imputation (Li, Yu, & Rubin, 2012). While there is no established cutoff for the proportion of data that can be missing, simulation studies have demonstrated multiple imputation techniques to be acceptable with 40% (Kristman, Manno, & Cote, 2005), 50% (Scheffer, 2002), and 60% missing data (Rubin & Schenker, 1986). Ensuring the appropriate strategy is applied given the mechanism of missingness is of greater

importance than the amount of missing data. However, given the large proportion of missing data, analyses were run using list-wise deletion and results were compared to multiple imputation. In situations where there were noticeable differences in effect size or significance, the relative efficiency of the imputed variable was examined and list-wise deletion would be used.

For the imputation procedure missing demographic and clinical measures were imputed using the Markov Chain Monte Carlo (MCMC) approach using SAS PROC MI to generate three data sets. Imputation models included the outcome variable, select auxiliary variables, and all variables used in the subsequent analysis including significant predictors of missingness. The imputed data sets were then analyzed separately and results combined with Rubin's rules using SAS PROC MIANALYZE (Rubin, 1987).

Data Analysis

The analytic process began with simple univariate and bivariate statistics to describe the sample population. These basic analyses are crucial to understand how constructs are documented and measured in order to inform more complex explanatory multivariate regression-type models. The data analysis plan for each specific aim and question will be described further below. All analyses were conducted in either SAS 9.3 or R 2.15.

Prior to reviewing the specific analytic techniques for each aim, a brief review of background information for three advanced analytic techniques will be described. Multilevel modeling, survival analysis and Cox regression, and propensity score analysis. These techniques are used in different research questions so a proper grounding in these methods is useful to provide sufficient background for the rationale of the methods and appropriate application to each research question.

Multilevel modeling. Multilevel modeling techniques provide an opportunity to further explore the contextual information that is inherent in the major theories describing child maltreatment. The ecological framework of child maltreatment describes precisely that this phenomenon is impacted by an expanding context of individual, family, community, and society across time. Our statistical analyses and analytic models should incorporate the structure implied in our theories (Coulton, Corbin, & Su, 1999). When group-level information is aggregated to the individual level, we have assumed that the relationships that have been modeled apply equally to all contexts. Also, the group-level information that has not been modeled is pooled in the error term. In the current example, modeling social factors that impact outcomes as individual variables would lead to families living in the same neighborhoods to have correlated error, a violation of one of the main assumptions of multiple regression, leading to biased estimates. Additionally, these limited models stand in contradiction to the ecological theory of maltreatment.

As a public health outcome with key social determinants, describing child maltreatment using multilevel modeling is consistent with the investigation of many environmental predictors, especially the effects of neighborhood on health and mental health outcomes (Diez-Roux, 2000; Luke, 2005). Research into the etiology of maltreatment has recently focused on not just individual and family predictors but also building models that explore the interactions between different levels of analysis (Coulton, Crampton, Irwin, Spilsbury, & Korbin, 2007). Specifically, researchers have begun to uncover how characteristics of neighborhoods impact the variable rates of maltreatment that exist. Although the association between poverty, social disorganization, stress, and maltreatment rates has been shown, few models have explored and tested the theorized explanations (Coulton, et al., 2007).

The data are structured across three levels of observation: the infant/caregiver, infant/caregivers clustered within nurses, and infant/caregivers clustered within zip codes. The time level of analysis is not used in this analysis as survival analysis models the time to the first event as opposed to measuring risk for report across different periods of time. Multilevel modeling was used to account for the interdependence of the identified observations by partitioning the variance across the individual, nurse, and community levels of data. This variation was described using a null model with no explanatory variables in order to examine the variance components of each level. The three nested levels of the null model can be specified as follows: child or caregiver (*i*), nurse (*j*), and zip code (*k*):

Probability of $\text{Report}_{ijk} = \text{logit}\,(\text{Y}_{ijk})$

Probability of $\text{Report}_{ijk} = \pi_{\text{o}} + f_{0k} + v_{\text{o}jk} + u_{0ijk} + e_{0ijk}$

$Var(\text{Y}_{ijk}) = \delta^2_f + \delta^2_v + \delta^2_u + \delta^2_e$

Intraclass correlation (ICC) was determined for the nurse and zip code levels to examine whether or not a random-effects model is necessary for each level and outcome of interest. The ICC for binary outcomes, such as initial engagement, were determined by examining the proportion of the variance component assuming the level-1 residuals follow the logistic distribution ($\pi^2 / 3$). Census 2000 information relating to community characteristics such as employment, poverty, education, and mobility were collected. Based on the agreement with the agency, no nurse-level covariates were available for this study. For data analysis, families that receive services from the same nurse or who live within the same zip code are likely to be more similar to one another, violating assumptions of independence in regression-based modeling. The nurse and zip code variables were entered as random effects into the statistical models to provide more unbiased estimates of treatment effects.

Survival analysis and cox regression. Survival analysis is a class of statistical techniques that involves the modeling of the amount of time it takes for an event to occur, or time to event data. While understanding how many reports of maltreatment or injuries occur between groups is important, it is also useful to measure the differences in the rate at which these events occur. Since events such as maltreatment have a non-normal risk distribution and are right-censored by the constraints of the study time period, traditional regression methods are not appropriate. Survival analysis involves estimating a hazard rate for subjects, which is the probability that an individual will experience an event during a given time while at risk for the event. Similarly, the survival function indicates the probability that a child in this sample will not be maltreated past a given time.

Bivariate survival analyses can incorporate the effect of baseline covariates on the hazard function. For example, the effect of treatment condition can be included to examine the impact on the hazard function. The Cox proportional-hazards regression model (Cox, 1972) describes a multiplicative relationship between multiple covariates and the hazard function at baseline and varying effects at other time points. The main assumption of the model is that the hazard functions for different levels of a covariate are the same at any time point. In bivariate survival analysis, covariates are limited to binary conditions while Cox models can be estimated with continuous covariates. Time-varying covariates can also be modeled using Cox regression to examine the change in effect of a covariate at different times.

Because families were at risk for a report for different periods of time following NFN services, survival analysis was used to analyze the rate of report occurrence over time. Exploratory univariate survival analyses were used to compare estimates of the survivor function (probability of not having a report at a specific time) across demographic variables, treatment

factors, clinical measures, and child welfare history. For categorical variables, Kaplan-Meier survival curves were created and the log-rank test of equality across strata was used to test significance. For continuous variables the Wald chi-square test from a Cox regression model with a single predictor was examined.

Multivariate Cox proportional hazards regression modeling, a form of multivariate survival analysis, was used to model risk for report controlling for covariates. Given that respondents are nested within a set of nurses, the survival times are assumed to be correlated. For this reason a robust sandwich covariance matrix estimate is used to account for the intracluster dependence in these models (Allison, 1995). The robust standard errors created from this step are generally smaller than those from a standard partial maximum likelihood estimate in Cox regression. For Cox regression models, the hazard ratio (HR) will be interpreted as opposed to the regression parameters. The HRs is similar to odds ratios and allow for easier interpretation of the findings. For each parameter, the HR is the ratio that a hazard rate changes for one unit increase in the covariate.

Propensity score matching. Aim 3 compared outcomes for treatment versus a matched comparison group created using propensity score matching (PSM) and other propensity score analytic techniques. The theoretical basis of this statistical technique is based in the counterfactual framework and the concept of exchangeability (Oakes & Kaufman, 2006). To truly examine a causal effect, one would have the impossible task of observing the outcome for an individual given an exposure and at the same time, observe what happened to the same individual without the exposure. Based on the statistical gold standard of the randomized experiment, the Nyman-Rubin counterfactual framework (Neyman, 1923; Rubin, 1974) posits that the true causal effect ("potential outcome") of treatment on the experimental group can be

estimated by comparing observed outcomes in a sample group not exposed to treatment (Guo & Fraser, 2010).

The assumption under these conditions is that randomization produces equality in the groups, or exchangeability. Thus, one control subject could be substituted for an experimental subject without penalty. The key assumption is that exchangeability "implies that an observed counterfactual substitute *could have been* treated or exposed (or both) just as the unobservable counterfactual could (theoretically) have been" (Oakes & Kaufman, 2006, p. 373). In this framework, exposure or treatment could have happened to anyone in the study, and a truly randomized experiment is able to obtain an unbiased estimate of the causal effects of the treatment (Rosenbaum & Rubin, 1983).

However, randomization is often impossible or impractical in social science, especially in the area of program evaluation (Guo & Fraser, 2010). For example, in the current study, there have been efforts to establish efficacy of nurse home-visiting program in randomized experiments. The next step is then to examine the effect of such programs when implemented naturally in a community environment, unconstrained by controlled trials, to established effectiveness. However, there is still a need to answer the key cause-and-effect question: "to what extent can the net difference observed in outcomes between treated and nontreated groups be attributed to the intervention, given all other things are held constant" (Guo & Fraser, p. 21-22). Thus, average treatment effects must still be estimated using observational data from quasi-experimental designs.

It is widely recognized that estimation of causal effects from non-experimental design using matching or statistical controlling can be flawed from sample selection bias (Guo, Barth, & Gibbons, 2006). New approaches to causal inference using observational data have emerged as a

class of statistical techniques known as propensity score analysis. The propensity score is simply the subject's probability of being treated "as a function of all relevant observed covariates-that is, observed pretreatment measurements possible related to posttreatment outcomes" (Rubin, 2010, p. 7). In a truly randomized study the propensity score for all subjects, treatment and control, is one half. There is a 50/50 chance that any given subject would be in treatment or control. In the case of NFN, one group consists of families who were referred to services and received at least two visits from the nurse. The other group consists of families who were referred and dropped out of treatment with only one visit. Clearly, the assignment to these two groups is biased due to the non-random nature of self-selection. There are many factors that would determine whether or not a caregiver would exist in the "treatment" group or would dropout to the "control" group.

The PSM model balances the treatment and control group through resampling and matching on the probability of being treated. In this case, NFN families who remained in treatment are statistically matched using the propensity score with those who dropped out to "mimic randomization" (Oakes & Kaufman, 2006, p. 376). Propensity scores are estimated using traditional logistic regression models, in this case the dichotomous outcome variable is whether the caregiver dropped out or stayed in treatment given a set of pretreatment covariates. For this analysis, the selection of covariates will include family demographic information, perinatal health information, maternal risk factors, and any other measured factors that might influence the selection into treatment or dropout groups. The groups are then balanced based on the propensity scores.

There are several methods of direct matching that can be attempted, but the 1:1 nearest neighbor within calipers is the most common and straightforward (Thoemmes & Kim 2010; Lanehart, et al., 2012). Propensity scores are probabilities that range between 0 and 1.0. Out of

the entire sample, one comparison subject is selected for each treatment subject based on the smallest distance between the PS. A specific maximum distance can be determined by the size of the calipers. Local or "greedy" matching algorithms make decisions at each step of the matching process for each individual unit without considering the best match that would minimize the overall difference (Coca-Perraillon, 2007). Global optimal matching algorithms minimize the overall distance between the propensity scores and the individual unit being matched. The adequacy of matching is assessed by minimizing the standardized differences between the two groups on the observed covariates.

The propensity score can balance only observed covariates, a major limitation of PSM and cause for concern when the matched group consists of dropout families (Joffe & Rosenbaum, 1999). While program dropouts can be problematic comparisons, they do meet the criteria of having similar initial levels of risk and represent a group that should be a reasonable test of program effectiveness. Data collected at baseline is substantial, lending confidence that a PSM approach will be successful. The NFN data are unusually detailed since case files are entirely electronic, which improved the counterfactual estimation and overlap in propensity scores between groups. Sensitivity analyses will also be conducted based on the matching technique selected and caliper width. Another benefit of this dataset is that families who drop out but remain in the sample can still be linked with administrative records of target child welfare and health information providing even more propensity covariates. The "MatchIt" package in R and the "vmatch" macro in SAS were used. There are several examples of this technique used in multilevel longitudinal and survival analysis (Haviland, Nagin, & Rosenbaum, 2007).

In addition to PSM techniques, the propensity score can be used in analysis other than matching to estimate treatment effects. Two other approaches are stratification and regression adjustment (D'Agostino, 1998). These methods will also be tested using the data to examine if the findings from the matching analysis are consistent. Since matching techniques generally require data reduction and loss of sample size since subjects without good matches are dropped, there can be some concern about generalizability. All three techniques estimate the propensity score using the same calculation. The difference is how the score is used. Using stratification, families in both conditions are ranked on their propensity score and categorized into equal sized subclasses (usually quintiles). Once there is acceptable balance on the covariates, subjects from treatment and dropout groups that are in the same stratum are considered part of a separate randomized experiment and compared directly (Rubin, 2009). Assuming perfect stratification, the treatment effect for each stratum are weighted and used to determine the overall average, unbiased treatment effect for the study. The choice of strata does influence the variance and bias of the final estimate with a tradeoff between size of the strata and potential bias (Myers & Louis, 2010). However, Rosenbaum & Rubin (1983) have suggested that about 90% of the initial bias due to selection can be eliminated by stratification.

Regression adjustment, the most commonly used propensity score method, simply incorporates the propensity score as a covariate in the multivariate models (Myers & Louis, 2010). Often, the propensity score is included in multivariate models with other selected covariates to adjust the estimate of the treatment effect. This has the same overall effect as including every variable that is used to develop the propensity score in the multivariate model. However, since the goal is to estimate propensity to treatment assignment, the propensity score is often estimated using complex interaction terms and quadratic functions first with no concern for

including too many variables in the model. Thus, this two-step process allows the investigator to include the propensity score and only the key variables of interest. Using a simpler model, one can focus more reliably on model fit characteristics based on the independent variables of concern. Often regression adjustment and stratification will be used in tandem as an efficient and unbiased estimator of the treatment effect (D'Agostino, 1998). This study tested the treatment effects for matching, stratification, regression adjustment, and combined methods.

There are only three other studies that could be identified in the literature that attempted to use propensity score analysis in non-randomized early childhood home visiting studies. Duggan and colleagues (2007) used PS analysis in the context of a larger randomized effectiveness study to compare families who received a high dose of service. Similarly, Hill, Brooks-Gunn, & Waldfogel, 2003 matched high-dose families with preterm, low-birth weight infants in the Infant Health and Development Program (IDHP) to assess treatment effect for child cognition to a control group. Finally, an evaluation of Early Head Start attempted to measure dosage effects but results were not reported as "approaches did not yield consistent, reliable results" (Love, et al., 2002, p. 90). The analytic approach for each aim and research question will be described in detail below.

Aim 1 data analysis. This aim is to describe service use patterns and explore whether dose of NFN service moderates outcomes. The research questions increase in complexity for exploratory to explanatory from bivariate to multivariate models.

Question 1.1: What proportion of families are initially engaged and retained in services? This first research question was primarily descriptive in nature. The purpose was to provide a summary of the service use patterns for the sample of families selected for this study. The goal of the analysis was to provide the agency a better understanding of the flow of clients

into and out of the program and also to compare the service usage of this group to other home visiting models. Beyond the application to understanding service utilization at the agency-level, this analysis will also help to better understand the service utilization of a nurse home-visiting program that has a flexible, nurse/client driven treatment plan. The level of engagement and retention can be compared to those models with prescribed numbers of visits. Bivariate analysis including chi-square, t-test, and analysis of variance (ANOVA) are used to compare families engaged to those who are not engaged. Given that the retention outcome (number of visits) does not have a normal distribution but a Poisson or count distribution, it does not lend itself to tests of association that assume variables are normally distributed. Therefore the retention variable was recoded as a categorical variable for prenatal visits (0, 1, 2-3, 4 or more), postpartum visits (0, 1, 2-3, 4 or more), and total home visits (1, 2, 3-4, 5-6, 7 or more).

Question 1.2: To what extent do engagement and retention vary by individual, provider, and neighborhood? Based on Daro & McCurdy's Theory of Parental Involvement (2001), there are factors across multiple levels of human ecology that are considered important in predicting the involvement of a caregiver in supportive services. As a voluntary service, this theory suggests that individual-level factors likely play the most important role in determining whether or not a caregiver will choose to become involved with services initially and then continue to participate in home visiting services over the recommended course of the program. The goal of this research question is to determine to what extent the child and family level, the nurse level, and the community level explain the variance in the outcomes of initial engagement and retention.

First, a predictive model using key child and family demographics characteristics and risk factors was developed predicting the initial engagement and retention. The variance explained

for these models can be determined using the R^2 model fit test statistic. This statistic ranges from zero to one and describes the percent of the overall variance in the outcome that can be explained by these individual level predictors. Families are clustered within nurses and as well as clustered within zip codes. Multilevel modeling techniques were used to estimate the intraclass correlation (ICC) at the higher levels as a measure of variance explained at the nurse and zip code level. While there is no accepted cut-point for a statistically significant ICC, typically an ICC over .05 is considered sufficient to warrant further modeling at that level.

Question 1.3: Which individual-level and zip-code level predictors are associated with engagement and retention? This research question examined specific predictors at the individual and zip-code level for engagement and retention. Nurse-level predictors were not available from the agency, so cannot be included in this analysis. If the ICC was over .05, zip-code level predictors would be used in the models. Zip-code level predictors were taken from the 2000 Census and were selected among those found to be correlated with service utilization and access.

A hierarchical multiple regression approach was taken for the model-building process. To account for clustering at the nurse level and non-normal outcome variables, a generalized linear mixed model was estimated using PROC GLIMMIX with a random intercept at the nurse level (Ying, 2006). Engagement and retention were regressed on blocks of covariates across four models. The first model contained child and caregiver demographic characteristics, the second model contained the demographics and individual risk factors, the third model contained demographic and service use variables, and the fourth model contained all variables. In the models with individual risk factors the cumulative risk score was removed to eliminate issues of

multicollinearity. For engagement, a binary outcome using the logit link was specified given the dichotomous engagement variable.

Count variables and rare events that take place in a given interval of time that take the form of a Poisson or negative binomial distribution which can be estimated with regression models expressing the natural logarithm of the outcome (Karazsia & van Dulmen, 2008). Outcomes with non-normal distributions violates the assumption of ordinary least squares regression and lead to biased estimates and poorly fitting models. In these models, an assessment of the deviance and residual plots was used to examine the assumptions of the model and the fit of the data. For retention (total visits), PROC GENMOD was first used to examine the dispersion characteristics of the visits count outcome. This was done to determine whether or not the variable took on a Poisson or negative binomial distribution. The Wald confidence interval for the dispersion parameter (.49 to .56) did not include 0, so the negative binomial distribution was appropriate the retention analysis.

Comparative model goodness-of-fit was assessed using the Akaike Information Criteria (AIC), with decreasing AIC suggesting an improvement in the model fit. For engagement, the point estimates and confident intervals were exponentiated to create odds ratios. These can be interpreted as an increase in the odds of engaging in services for each unit increase in the given covariate. Odds ratios greater than significantly higher than one are associated with higher odds of engagement while ORs below one are associated with lower odds of engagement. For retention, beta coefficients can be interpreted as the expected change in the log count of visits. Betas significantly greater than zero are associated with a higher number of visits while betas below zero are associated with fewer visits.

Question 1.4: Is retention associated with greater documented skills training? This analysis explored whether or not more visits was associated with greater acquisition of skills as documented by the nurses. A fairly straightforward assumption of the program is that a higher dose of home visits ultimately leads to a higher number of skills that can be imparted from the nurse to the caregiver. The home visiting conceptual model assumes that during the home visits there is a change in knowledge, behavior, and attitudes of the caregiver that plays as a mediating role in impacting the distal outcomes. Given that visits last approximately one to two hours, it is not possible to teach the caregiver all the required topics in one visit. However, with attrition from services a common concern, it is important to understand what depth and coverage of topics can be covered and what is the added value from an additional visit, on average. There is an assumption that more visits equates to more skills, however, it may be the case that some caregivers require more visits and contact hours to get to the same level of information. This analysis first explored the bivariate relationship between visits and skills. Preliminary analysis determined that the distribution of skills took on a relatively normal distribution (M = 24.6, SD = 11.2, skewness = .19) allowing for multiple linear regression to be used. The count of total skills was regressed on root-transformed total visits and a set of control variables to determine the relationship between visits and skills controlling for other factors.

Q1.5: Are services associated with risk for maltreatment? *If so, is there a dose response relationship?* This research question addressed the extent to which visits moderated risk for maltreatment for the entire sample or for specific covariates. Moderation analysis is accomplished by entering an interaction term in addition to the main effects for the variables in the models. If the interaction term is significant this means that the main effect relationship between a given covariate and the outcome is dependent on the number of visits that are

received. For a significant interaction term that is negative, this can be interpreted that a higher number of visits decreases the main effect hazard rate for that covariate. For this analysis a categorical visits category was entered in the model (1 visit, 2-3 visits, 4-7 visits, >7 visits) as an interaction term with child and caregiver demographic characteristics and risk factors. This allows the analysis to consider different levels of dosage as a moderator of outcomes. Specific covariates of interest were tested individually including moderation by level of cumulative risk and across program service populations. Then an exploratory model including child and caregiver demographics and risk factors was estimated. Interaction terms that were significant at $p < .10$ were considered at the bivariate level for evidence of trends.

Aim 2 data analysis. To compare maltreatment reports, child development, and maternal mental health outcomes among those served in key policy-relevant subgroups.

Question 2.1: Are there differences in rates of child maltreatment reports for family subgroups? This analysis explored the difference in the rate of maltreatment reports during the two-year follow-up period for the identified subgroups. There is a large body of literature that connects the risk for maltreatment to different subgroups of families.

Question 2.2: Are there differences in maternal mental health indicators for family subgroups? Given the importance of maternal mental health outcomes as a mediating factor in risk for child maltreatment as well as engagement in services, this analysis examined the levels of stress and postpartum depression across the identified subgroups based on screening tools deployed by the nurses during the first visit. Stress was determined using the Everyday Stress Index and depression using the Edinburgh Postpartum Depression Screening tool. Levels of self-reported mental health diagnosis were also examined based on the initial nurse assessment.

Question 2.3: Are there differences in child developmental health indicators for family subgroups? The ASQ was used by nurses to screen for developmental concerns for the infants served by the program. Research in this area suggests that risk for developmental problems vary across risk subgroups.

Aim 3 data analysis. A quasi-experimental design was selected in part because the program is already "saturated" in the service areas and it is therefore not feasible or ethical to randomize families in need of services to a "no services" or waitlist condition. Using advanced propensity score matching techniques (reviewed above), families who drop out of the NFN program after having completed the initial assessment process were compared to families who complete the program to assess effectiveness of the intervention. Existing electronic case data collected by the agency were linked to data from state administrative records for child health and maltreatment outcomes.

The framing of the research questions for this aim are purposefully worded to reflect the language typically used in the causal analysis methodology literature (Guo & Fraser, 2010). As the primary aim of this study, it reflects the primary mission of this agency and evidence-based home visiting broadly. In theory, the services provided by the agency purport to "cause" a decrease in risk for child maltreatment. Caution is always used when attempting to establish causality in the social sciences. The research questions will first be introduced then a rationale for the methods used will be provided describing the theory behind estimating a causal treatment effect using observational data from quasi-experimental designs. In an ideal situation, families would be randomized to a condition of NFN and others to a condition of no NFN, services as usual or a waitlist control group. Since this is not the case for this study, attempts to measure treatment effects are not simply abandoned but estimated using advanced statistical methods.

Q3.1: To what extent can any differences in the rate of maltreatment reports be attributed to the Nurses for Newborns intervention? This question examines the raw percentage of families who experience a maltreatment report at any point during the follow-up period. The raw report rate for all families is a useful value to obtain, however to determine whether the program had a causal impact, we must whether or not this value is significantly higher or lower than would be expected from a similar comparison group. The propensity score techniques described below attempt to remove a potential selection bias from those who remain in the program in order to ensure that any differences are due to involvement in the program, not to other factors that might be common to families who stay in treatment and who also have a higher risk for maltreatment.

Q3.2: To what extent can any differences in the timing of maltreatment reports be attributed to the Nurses for Newborns intervention? This question is a continuation of the previous question but also adds a time element to the analysis. It is possible, for example, that the program has an effect that only lasts for a very short term after the family is exposed to the intervention. The use of advanced survival analysis techniques compares the time to event, in this case a maltreatment report, among the two groups of interest.

Client Flow

The study sample was divided in two groups at several points in the analysis between families who enrolled prenatally and those who enrolled postpartum. This was due to the fact that there was not available birth data on a large number of families who enrolled prenatally and dropped out before the agency was able to collect information about the birth of the child, including identifying information. This information (birth date, Social Security, and name) was used to match with state CPS records. For reasons of data availability and potential bias

associated with those families for whom infant information could not be collected, the later CA/N outcomes analysis was limited to those with birth information and only for those who had at least one postpartum visit. However, there was a similar level of data collected for prenatal mothers regarding risk factors, engagement, and maternal mental health outcomes. The following Figure 1 presents a client flow diagram to describe how many families were excluded at which time point for prenatal and postpartum families. This Figure also includes the rate of later reports at the bottom across the treatment (>1 visit) and comparison (1 visit) conditions after sample exclusion. Not that the samples are labeled A (total sample prior to exclusion), B (limited to those with birth data), and C (those with birth data and at least one postpartum visit).

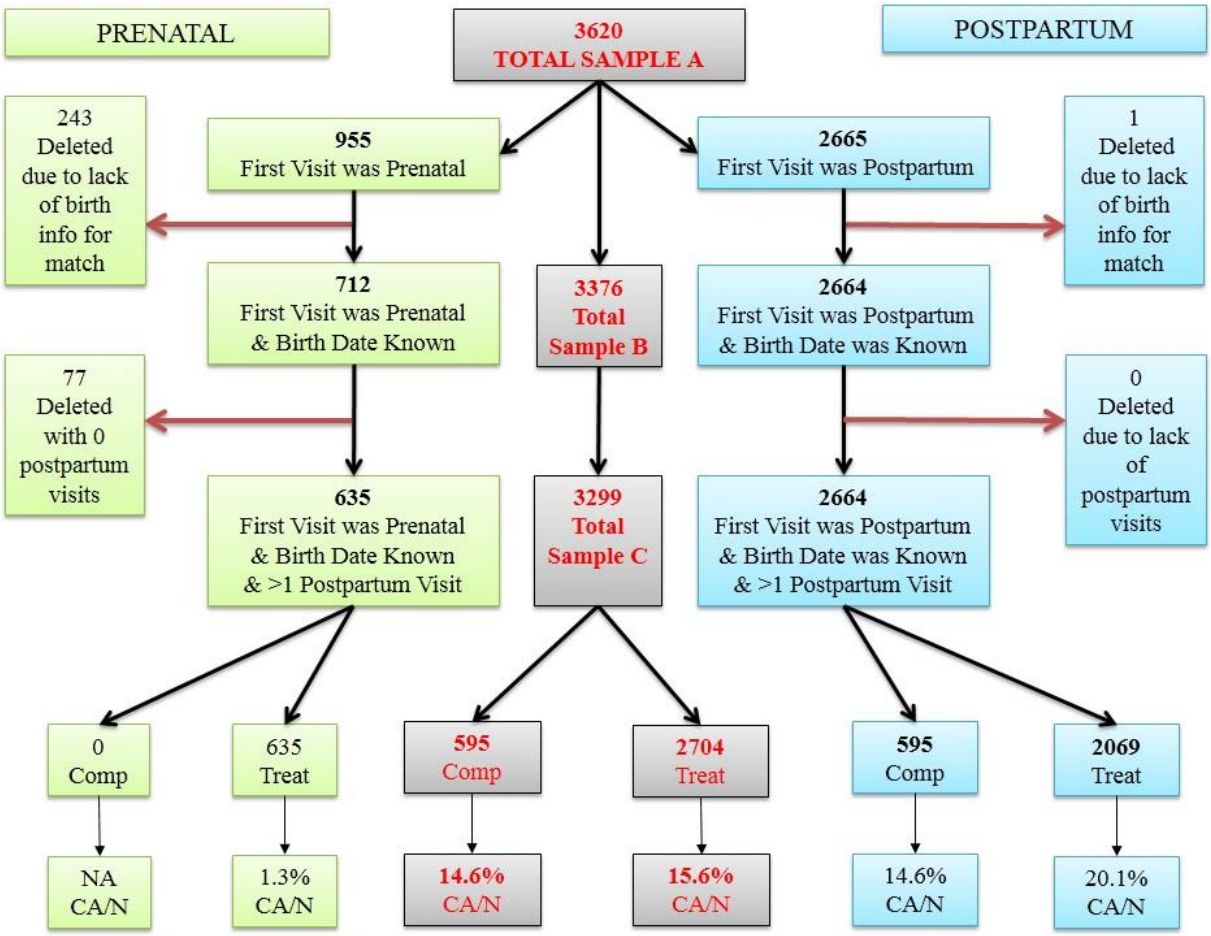


Figure 1. Client flow diagram. This figure provides further description of the number of families who were excluded due to missing infant birth data and those who lacked at least one postpartum visit from the prenatally referred group. Rates of later CA/N for those in the comparison and treatment groups are also reported.

Chapter 4: Results

The results section will be organized by research question and each specific aim. The first aim addressed service use among participants in the sample and begins with a description of the service population. The bivariate descriptive results are organized by caregiver and child demographic characteristics (Table 2), risk factors (Table 3), and service utilization variables (Table 4). The first bivariate contrast across these variables is the relationship between the given characteristic and the prevalence in the four subpopulations. The primary purpose for displaying the results in this way is to examine the distribution of these characteristics across the populations as defined by the agency in order to determine where similarities and differences exist. This helps to better characterize the risk profiles of these groups in later analyses where differences in outcomes may exist. Instead of using the agency terms, the tables use a more descriptive label for the subgroups: primary poverty (Bright Futures), high-risk baby (Bridge to the Future), high-risk caregiver (Safe Beginnings), and teen parent (Teen Parent).

Results of Aim 1: Service Utilization

Aim 1: To describe service use patterns and explore whether dose moderates outcomes.

The maternal and child characteristics for the service population included in the sampling frame are provided in Table 2 below. The table contains infant characteristics including child age at follow-up (December 2012), race, ethnicity, gender, low birthweight status, and preterm birth status. The caregiver characteristics include relationship to the infant (biological mother or not), age at referral, urbanicity, prior pregnancy, other living children, marital status, level of education, and employment status. The maternal and child characteristics have significantly different representations across the service populations except for infant ethnicity and gender.

The demographic makeup of these different service group are notable, however, there is still a mix of different types of families in all of the groups.

Table 2

Demographic Characteristics of the NFN Program and Service Subpopulations

	Total Sample (n=3,620)	Primary Poverty (n=1,390)	High-Risk Baby (n=852)	High-Risk Caregiver (n=776)	Teen Parent (n=602)	
Infant Characteristics	100.0	38.4	23.5	21.4	16.6	sig.
Child Age at Follow-up	2.9 (.8)	2.9 (.7)	2.9 (.9)	2.9 (.7)	2.8 (.7)	*
Infant Race						*
Black	46.3	38.2	26.7	15.0	20.0	
White	46.2	35.4	23.2	28.4	13.0	
Biracial	6.7	37.8	23.1	25.8	13.3	
Infant Ethnicity (Hispanic)	3.7	39.7	21.4	20.6	18.3	ns
Infant Gender (female)	48.4	37.5	25.8	21.2	15.6	ns
Low Birthweight	17.2	8.5	69.3	19.2	3.0	*
Very Low Birthweight	2.5	3.6	83.3	10.7	2.4	*
Preterm Birth	18.6	11.6	66.0	20.4	2.1	*
Caregiver Characteristics						
Relationship to Infant (bio mother)	96.9	39.1	22.4	21.4	17.2	*
Age at Referral (<20 years)	36.1	19.5	19.1	15.4	46.0	*
Zip level Urbanicity						*
Urban	78.1	39.6	24.0	19.6	16.8	
Rural	13.8	31.1	23.3	31.7	14.0	
Suburban	8.1	40.2	18.9	21.7	19.2	
Previous Pregnancy	52.7	42.7	27.2	25.3	4.8	*
Living Children	41.3	40.5	30.0	26.7	2.7	*
Marital Status						*
Single	78.9	38.1	21.8	19.6	20.5	
Married/Consensual Union	17.4	42.0	30.4	25.6	1.9	
Separated	1.7	32.3	27.4	37.1	3.2	
Divorced	1.7	35.6	25.4	39.0	0.0	
Level of Education						*
No HS	37.3	24.7	21.5	23.8	30.1	
HS/GED	37.3	41.7	23.7	21.1	13.5	
Some College	19.9	55.7	24.1	18.3	1.8	
College Degree	5.6	49.0	30.8	19.7	0.5	
Employment Status						*
Unemployed	50.6	40.9	22.6	24.4	12.1	
Disabled	2.0	11.3	14.1	74.7	0.0	
Homemaker	8.0	41.3	31.8	23.4	3.5	
Student	15.9	15.9	16.7	12.7	54.8	
Part-Time	10.8	45.9	25.4	18.1	10.6	
Full-Time	12.5	54.9	28.8	13.6	2.7	

Note: *$p < .05$

The first distinction that is made is in the teen parent service population, which makes up 16.6% of the entire sample. While 36.1% of the sample is below the age of 20, only 46% are in the teen parent subgroup. NFN assigns families to service populations in a hierarchy where having a medically fragile newborn or a significant mental health issue will "trump" the age of the mother. This is one example of the multi-risk nature of the sample. Families do not easily fit into one specific category of risk. Overall though, caregivers in the teen parent population were more likely to be on their first pregnancy, be single, have not completed a high school education (but to currently be a student), and live in a suburban area.

The high-risk baby (23.5% of the sample) and high-risk caregiver (21.4%) service populations also have significant trends in the demographic of the mother and child, but the distinction is not completely uniform. The infant in the high-risk baby group is more likely to be African-American, low-birth weight and preterm, and the caregiver is more likely to have a partner, have greater education, and more likely to employed. The high-risk mom group is more likely to be White, rural, separated or divorced, have lower education, and be unemployed or disabled. This distinction highlights that in this home visiting population, the risk can be distributed differently across the caregiver and the child, while in some programs these families would be excluded from services. The level of risk in the mother and child is different depending on the individual family and the population served.

The next table displays the presence of caregiver and child risk factors across the different service populations. The caregiver risk factors are provided across psychosocial, violence exposure, behavioral health, and maternal health outcomes. Overall, the total sample population is a very high risk group of families. All families in the sample were receiving Medicaid, WIC, or Food Stamps so the risk factors provided here are in addition to a generally

lower socioeconomic status and higher utilization of public assistance. The infants in this sample are also higher risk. In the general population, about 8% of children are born low birthweight and 12% are born preterm. In this sample, 17.2% are born low birthweight, 18.6% born preterm, and 23.7% were born with either status. In the at-risk baby service group, 63.9% are preterm or low birthweight.

Similar to the demographic characteristics, the risk factors are not completely isolated in specific service populations. However, the association between each risk factor and service population is significant at $p < .01$, so there are some clear differences in the risk profiles of the service groups. The families have a large number of risk factors across different domains. Overall, the high-risk caregiver service group has a significantly higher number of cumulative risks compared to the other service groups.

Table 3
Caregiver and Child Risk Factors across Service Populations

	Total Sample (n=3,620)	Primary Poverty (n=1,390)	High-Risk Baby (n=852)	High-Risk Caregiver (n=776)	Teen Parent (n=602)
Caregiver Risk Factors	100.0	38.4	23.5	21.4	16.6
Psychosocial					
Homeless	5.4	50.3	11.3	24.6	13.9
Father Unknown/Not Involved	2.5	33.7	16.9	31.5	18.0
Prior DFS Involvement	7.1	24.6	32.8	37.9	4.7
Teenage Mother	36.1	19.5	19.1	15.4	46.0
No High School Education	36.6	24.7	21.5	23.8	30.1
Multiple Current Children	41.3	40.5	30.0	26.7	2.7
Unemployed	49.9	40.9	22.6	24.4	12.1
Violence Exposure					
Current Domestic Violence	5.2	35.3	16.0	42.3	6.4
History of Rape/Sexual Abuse	3.6	36.2	8.5	46.2	9.2
History of Physical Abuse/Neglect	4.4	33.1	9.4	46.3	11.3
Behavioral Health					
Alcohol Use During Pregnancy	4.6	26.1	25.5	38.8	9.7
Drug Use During Pregnancy	13.3	17.5	46.0	31.4	5.2

Smoking During Pregnancy/In Home	24.4	33.0	25.2	30.3	11.4
Mental Health Diagnosis	20.0	10.8	6.5	80.1	2.6
Cognitive/Learning Disabilities	4.6	5.5	2.4	91.5	0.6
Neurological Impairments/Injury	2.0	11.3	14.1	69.0	5.6
Maternal Health					
Intended Pregnancy	14.6	42.7	26.4	24.1	6.7
Prior Low Birthweight/Preterm	3.2	33.3	33.3	30.8	2.6
Prior Fetal Death	8.4	44.9	20.5	31.0	3.6
Prior Pregnancy Complication	3.2	47.9	20.5	26.5	5.1
Current Pregnancy Complication	10.2	33.6	32.3	21.1	13.0
Inadequate Prenatal Care	21.6	29.3	30.8	21.2	18.7
<18 Months Between Pregnancies	9.4	47.5	20.9	26.0	5.6
Chronic Physical Health Problem	14.5	31.3	26.2	33.8	8.8
Child Risk Factors					
Special Care/NICU	4.0	6.2	71.9	19.9	2.1
Low Birthweight/Preterm	23.7	11.6	63.9	21.1	3.4
Drug-Exposed	7.6	1.8	69.8	28.0	0.4
Jaundice	2.8	23.0	41.0	22.0	14.0
Heart/Lung Complications	6.1	6.8	63.1	23.9	6.3
Major Congenital Disability	2.2	7.5	67.5	22.5	2.5
Prior CA/N Reported Victim	2.0	27.0	39.2	27.0	6.8
Cumulative Risk Score	3.9 (2.3)	3.0 (2.0)	4.8 (2.2)	5.4 (2.4)	3.2 (1.5)

Note: All risk factors and risk score significantly different across service groups at $p < .01$

Q1.1: What proportion of families receives a second visit (initial engagement) and what is the distribution of number of visits, hours of service contact, and weeks of agency contact (retention)? The following survival curves (Figure 2 and Figure 3) show the percentage of families still receiving services across the first ten visits. The figures are separated by those who receive at least one postpartum and only prenatal visits. Among all families, 79.3% receive a second visit, regardless of whether those services are initiated in the prenatal or postpartum period. By timing of initial visit, 51.1% of mothers that received only prenatal visits have a second visit, 77.7% of mothers initiating service in the postpartum period had a second visit, and 100% of those receiving both prenatal and postpartum (by definition) have a second visit.

The line at the 50% mark illustrates at which point half of the families that had received a first visit had been terminated from services. There are two distinct groups among those who engage prenatally. Although technically a mother enrolling in the prenatal period but discontinuing services prior to the birth could have multiple visits, a subgroup tends to discontinue services rapidly; 50% terminated after two visits. If a mother initiated NFN prior to the birth but decided to continue services after the child was born, they were more likely to remain engaged. About half of the prenatal plus postpartum group is still receiving services at eight visits.

The median number of visits for the group that began in the postpartum period was about four visits. From this figure it appears that there is some variation in the level of service utilization depending on what point in the pregnancy and postpartum period services begin and whether or not there is consistency from before to after the baby is born. Those families who initiate prior to the child's birth and continue to have visits after, had more visits overall than those who initiated services after the child's birth.

The next figure displays the number of weeks that a family is engaged with services from the first to the last visit according to the service group. This includes families that drop out after one visit, so their service contact would be less than one week in all cases. There is a wide variation in service utilization from less than one week to over 16 weeks. The high-risk baby and primary poverty groups tend to have less overall time in the program, while the high-risk caregiver group tends to remain in contact with NFN the longest; 41.5% stayed more than 16 weeks.

Figure 2: Survival Curve for first 10 Visits by Service Receipt for Caregivers with at least 1 Postpartum Visit

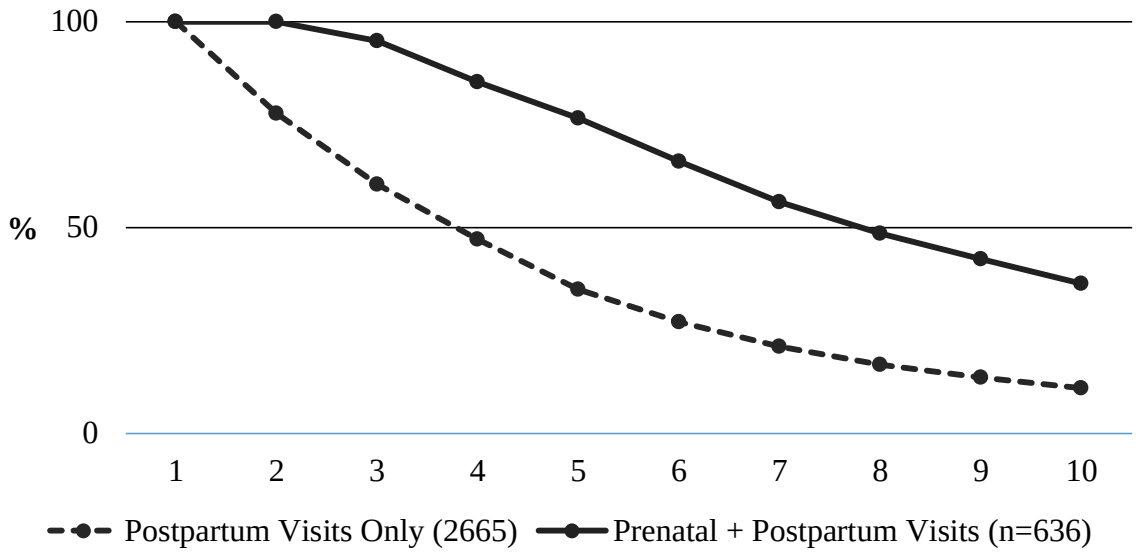
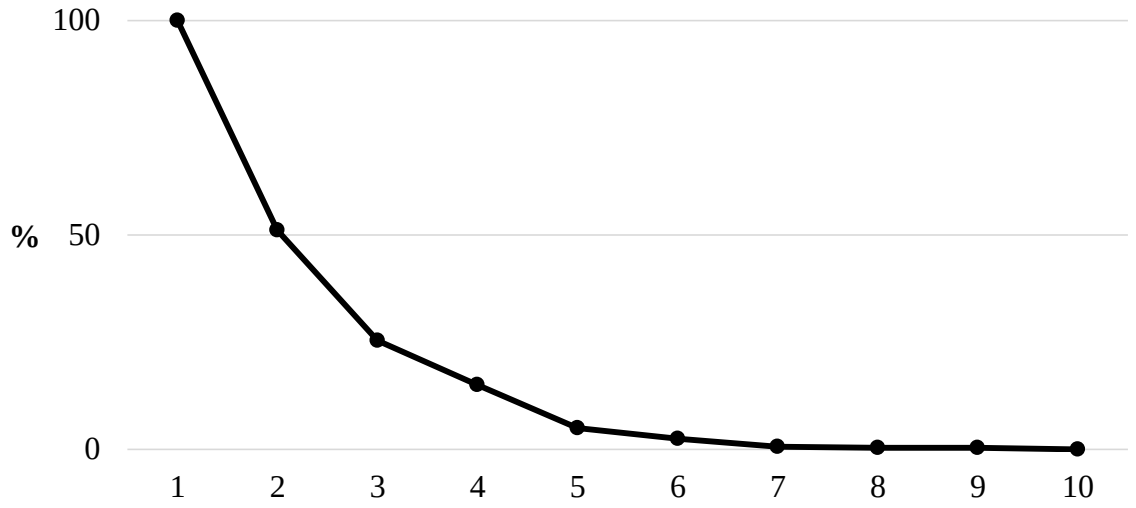


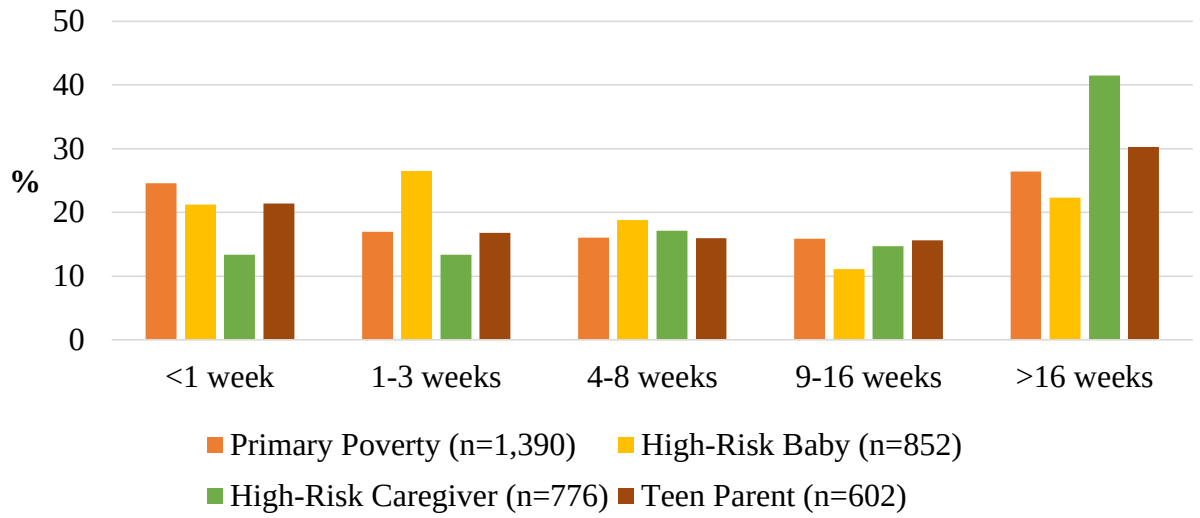
Figure 3. Survival Curve for first 10 Visits for Caregivers with Prenatal Visits Only (n=319)



While most visits last about an hour, the length of a given visit can vary. The next figure shows the number of hours of direct service contact that families receive across the four service populations. The vast majority of families received less than 10 hours of contact with the nurse. Similar to the distribution of weeks of service, the high-risk caregiver group is more likely to receive more hours of service contact, with 10.9% receiving more than 20 hours of nurse home

visit contact. By contrasting this with the prior figure, one sees that although the high-risk caregivers tend to remain on the active caseload longer, this does not necessarily translated into a significantly higher number of actual contact hours. This may occur due to missed appointments or difficulty scheduling visits.

Figure 4. Distribution of Number of Weeks from First to Last Visit by Service Subpopulation



The next bivariate descriptive table provides an analysis of association between several service utilization characteristics and the four service subgroups. This table describes at what point in the pregnancy or postpartum period the referral is made, when services are initiated, and the number of visits that are received in the prenatal and postpartum period. The table also reports the referral source and termination reason. All relationships are tested using chi-square analysis. The high-risk baby group is more likely to be referred after the birth of the child while other groups are more likely to be referred prenatally. The vast majority of self-referrals are primary poverty families and they tended to have higher representation among caregivers who waited longer to have their first visit. Similarly, the high-risk baby group is more likely to receive no prenatal visits but more postpartum visits. Overall, the high-risk caregiver group

receives more total visits. The most likely groups to receive a termination code indicating program completion were the primary poverty and high-risk caregiver families.

An attempt was made to use termination codes as a proxy for program completion, but nurse documentation was not always clear. In addition to the termination codes, nurse clinical documentation at the final visit was analyzed to further identify details surrounding the case closure, although again the level of detail and consistency in notation did not always make for clear distinction. Roughly 51% of families passively refused services (simply stopped keeping appointments or returning calls) and another 18% actively refuse services (directly indicated to the nurse that they no longer wanted to receive visits). Based on these notes, about 10% of families have their cases closed because the nurse indicates that the family has fully benefitted from the program and no longer needed services. Another 5% close because of duplication of services and about 10% moved out of the service area during services. It is clear that in the majority of cases, families slowly fall out of contact with the nurse through missed visits and ignored attempts by the nurse to contact the family.

Figure 5. Distribution of Hours of Service Contact by Service Subpopulation

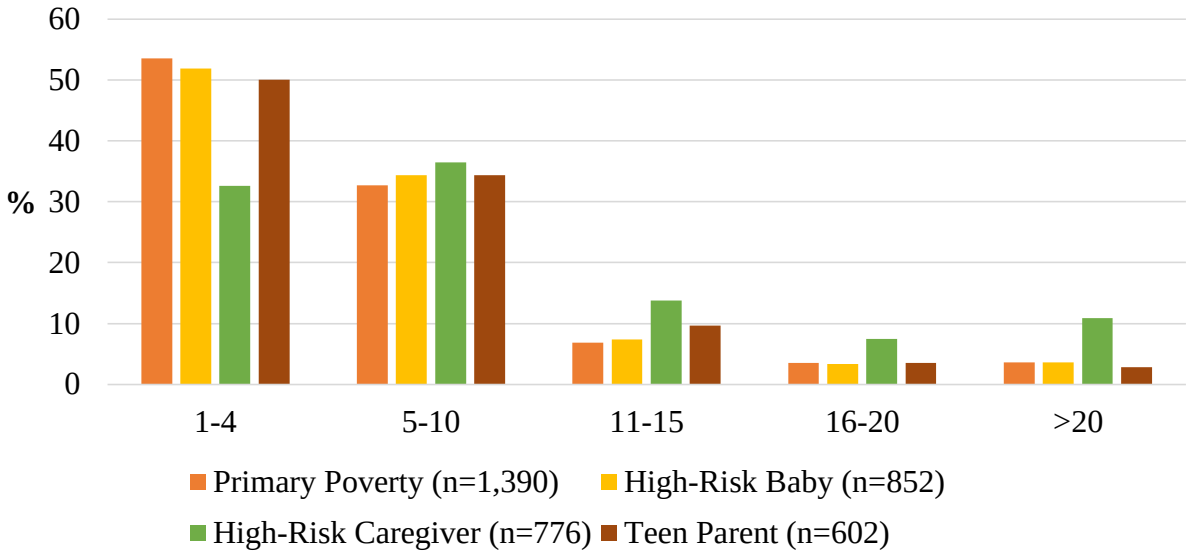


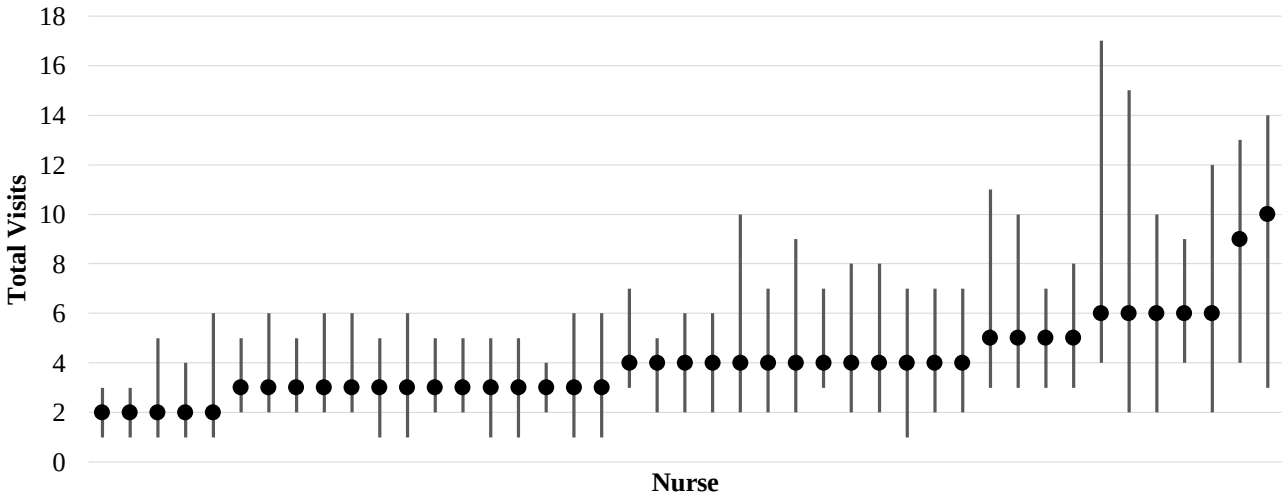
Table 4
Service Utilization Characteristics of the NFN Program and Subpopulations

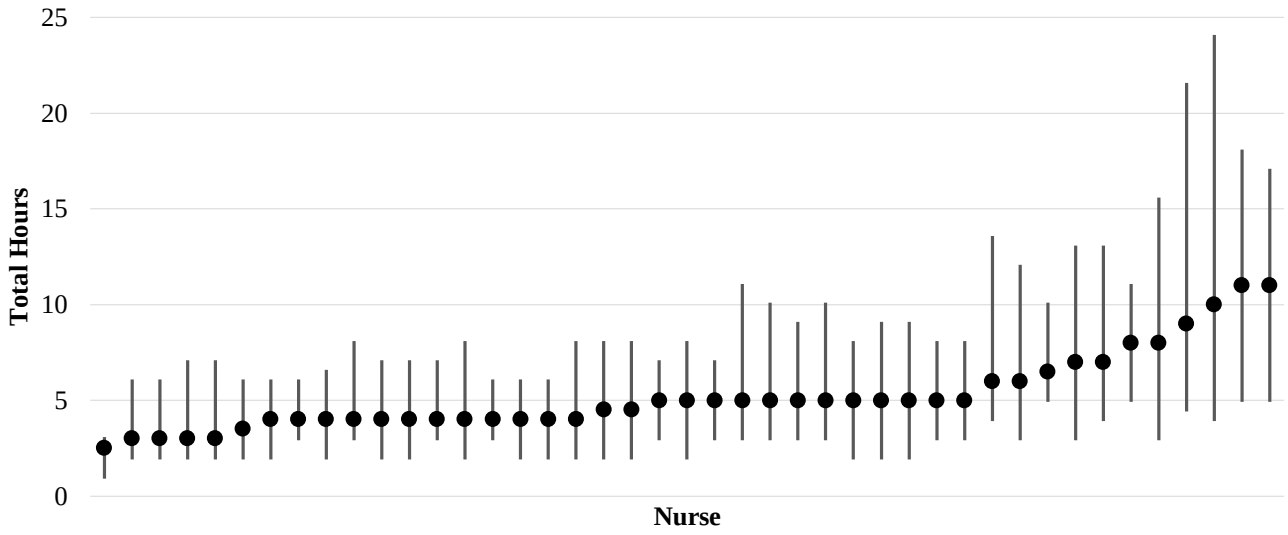
	Total Sample (n=3,620)	Primary Poverty (n=1,390)	High-Risk Baby (n=852)	High-Risk Caregiver (n=776)	Teen Parent (n=602)	sig.
Percent of Sample	100.0	38.4	23.5	21.4	16.6	
Infant Age at Referral						*
Prenatal > 90 days birth	19.8	46.4	5.9	20.4	27.3	
Prenatal < 90 days before birth	11.9	49.2	9.3	21.6	20.0	
Postpartum <30 days after birth	55.9	35.3	28.2	21.7	14.8	
Postpartum 30 - 90 days after birth	5.9	25.2	46.3	23.4	5.1	
Postpartum > 90 days after birth	6.6	32.9	42.6	20.3	4.2	
Referral Source						*
Self-Referral	18.3	59.0	12.8	17.9	10.2	
Healthcare Agency	57.5	30.9	28.6	19.9	20.6	
Social Service Agency	23.8	40.8	19.5	27.4	12.3	
Time from Referral to First Visit						*
1 day	9.9	36.1	22.1	30.3	11.5	
2-6 days	37.2	36.8	25.3	22.0	15.9	
7-30 days	43.2	40.5	23.4	20.2	15.9	
>30 days	9.8	37.8	18.6	15.8	27.9	
Prenatal Nurse Visits						*
0 Visits	73.6	34.2	30.0	21.4	14.4	
1 Visit	9.7	54.6	7.1	18.5	19.9	
2-3 Visits	9.8	50.1	4.8	19.0	26.1	
4+ Visits	6.9	44.0	4.4	29.2	22.4	
Postpartum Nurse Visits						*
0 Visits	8.8	58.0	1.6	16.3	24.1	
1 Visit	18.8	42.6	27.0	14.3	16.2	
2-3 Visits	27.2	39.5	26.9	16.3	17.3	
4+ Visits	45.3	32.2	24.4	28.5	15.0	
Total Nurse Visits						*
1 Visit	20.7	45.5	23.6	13.9	17.1	
2 Visits	15.8	44.1	25.7	14.2	16.1	
3-4 Visits	23.8	36.6	25.4	20.2	17.8	
5-6 Visits	14.3	36.6	24.6	22.7	16.1	
7+ Visits	25.4	31.8	19.8	32.6	15.8	
Termination Reason						*
Refused Services	31.7	37.7	22.7	21.8	17.8	
Moved/Cannot Locate	7.0	44.5	16.2	23.1	16.2	
No Need/Duplicate Services	2.5	30.3	33.7	19.1	16.9	
Completed 24 months	1.0	43.2	13.5	37.8	5.4	
Other Reason	57.7	38.9	24.3	20.7	16.2	

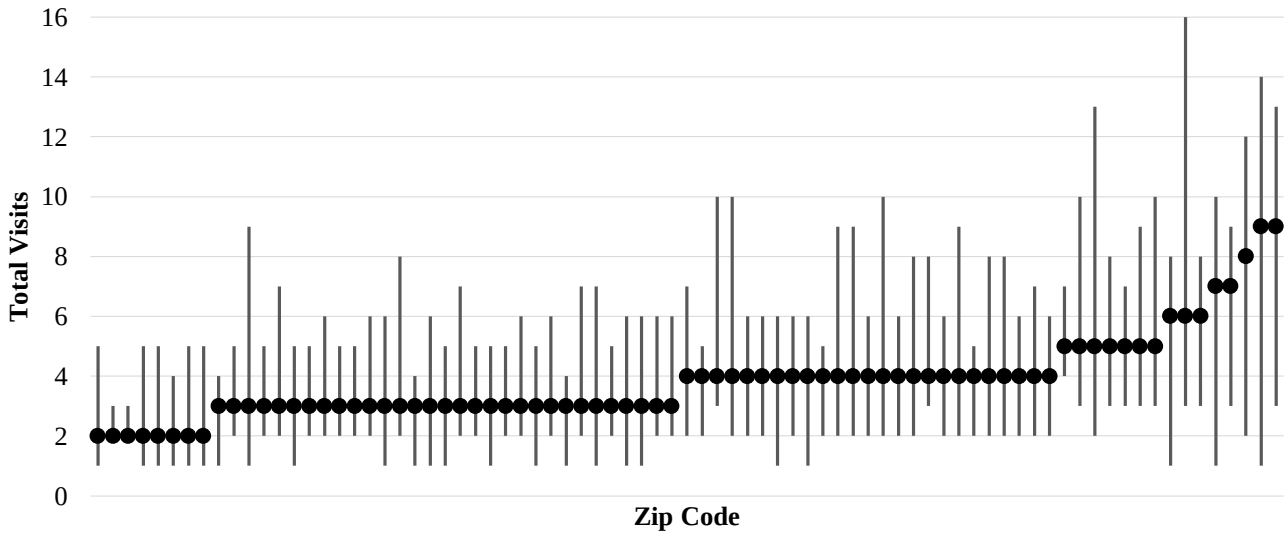
Note: *$p < .05$

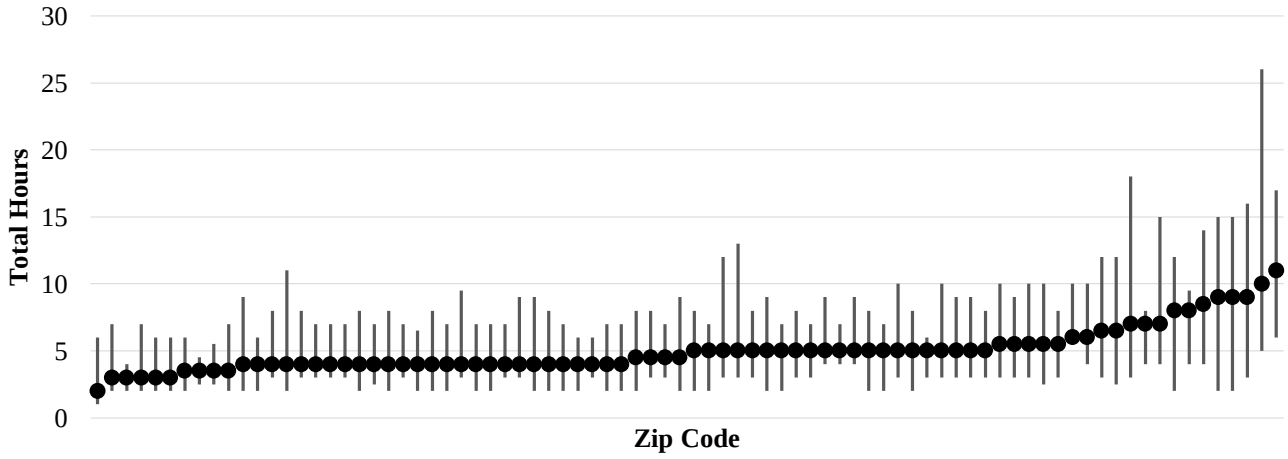
Q1.2: To what extent do engagement and retention vary by individual, provider, and neighborhood? This research question is concerned with describing the amount of variance in the engagement and retention outcomes that can be explained at the three levels of analysis predicted by theory. One way to examine this variation is to explore outcomes across the different levels of analysis. The previous question explored engagement and retention at the individual level and found considerable variation in the number of visits received by individuals in the service population. In subsequent models predicting engagement and retention, the R^2 or pseudo-R^2 for models predicting engagement and retention were between 7-16% depending on the individual parameters entered in the model. This suggests that a good portion of the variance in these outcomes can be explained by infant and caregiver factors. The following set of descriptive plots (Figure 6) provide the median, first quartile, and third quartile for the number of visits and number of hours across nurse and zip code clustering. These plots are limited to those nurses who served at least 20 families and those zip codes with at least 10 families. The plots are sorted by ascending median to give an idea of the range of median visits and hours. The vertical bars display the interquartile range for each nurse or zip code for each retention outcome.

Figure 6: Variation in visits and hours of contact by nurse and zip code clustering




25
20
15
10
5
0
Total Hours
Nurse


16
14
12
10
8
6
4
2
0
Total Visits
Zip Code


30
25
20
15
10
5
0
Total Hours
Zip Code

After these descriptive visual plots are used to display the variation in the outcome of interest, formal analyses can be conducted to determine the share of variance across the two higher levels of clustering. Using multilevel modeling techniques, the intraclass correlation (ICC) was determined for the nurse and zip code level. This was calculated by fitting a null model and examining the variance components first with a two-level including the individual and either the nurse or the zip code. Next, a three level model was fit including both nurse and zip code. In the model predicting number of visits, the ICC for nurse-level was .114 and the ICC for zip-level was .008; this corresponds to 11.4% and 1.0% increase in the variance explained by including this level of analysis. The ICC for number of days of service contact was 10.0% and 0.7% and the ICC for number of hours 13.7% and 0.1% for nurse and zip level respectively. The model fit predicting each outcome was examined and was improved by adding the nurse level but remained the same when adding the zip level random effect. The ICC predicting CA/N outcome was also examined at this point and based on the low ICCs for both nurse and zip, the nurse clustering level was used for all analyses in this study.

Table 5
Intraclass Correlation (ICC) for Null Multilevel Models Predicting Engagement, Retention, and CA/N Report

Service or Outcome Measure	Nurse	Zip Code
2-Level Initial Engagement	0.051	0.015
3-Level Initial Engagement	0.050	0.004
2-Level Visits	0.164	0.077
3-Level Visits	0.114	0.008
2-Level Hours	0.198	0.109
3-Level Hours	0.137	0.007
2-Level Days	0.089	0.029
3-Level Days	0.100	0.007
2-Level CA/N Report	0.030	0.035
3-Level CA/N Report	0.018	0.012

Q1.3: Which individual-level predictors are associated with engagement and retention? Engagement and retention outcomes were regressed on individual-level predictors using a hierarchical regression model entering covariates in blocks. The first analysis predicted the odds of the binary outcome of engagement estimated using a generalized linear model including an intercept-only random effect at the nurse level. The second analysis predicted the number of visits using a negative binomial generalized linear model. The results of these models are shown in Table 6 and Table 7. Only significant covariates are included in the tables. Model 1 included caregiver and infant demographic characteristics and all significant predictors carried through to subsequent models, so Model 1 was not shown for ease of interpretation of findings. The reporting of results will focus on the final full model.

The final engagement model (Table 6) provides several strong predictors of initial engagement based on specific risk factors as well as service utilization patterns. Significant risk factors predicting initial engagement include caregiver history of CA/N, cognitive or learning disabilities, prior fetal death, and infant low birth weight or premature status. Caregiver history of maltreatment and/ or presence of cognitive delay had the strongest association with engagement (both increased likelihood over 2.5 times). Factors relating to lower odds of initial engagement were unintended pregnancy and caregiver chronic physical health problems. Families referred during the prenatal period were more likely to engage with services than families referred later after the child was born. Social service agency referrals were more likely to engage than self-referrals. Families that had their first visit within a day of the initial referral were nearly 2.5 times more likely to engage in services.

The model predicting number of visits (Table 7) indicated that there were some unique factors that predicting ongoing participation as compared to initial engagement. The parameters

for these models can be interpreted as the change in the log-odds of the number of visits for each unit increase in the given covariate. Similar to the engagement model, many of the caregiver and child risk factors predicted a greater number of total visits. While cognitive delay and a history of maltreatment were significant predictors of both initial engagement and number of visits, the presence of domestic violence was only significant in the model of number of visits (increasing number of visits). Low birthweight babies and those with heart or lung issues also had a higher number of predicted visits. Similar to engagement, families who were referred earlier in pregnancy and had less delay between referral and the first visit had a higher number of visits.

There were some similarities and differences in family characteristics between initial engagement and number of visits. While race/ethnicity was not significant in the final model for initial engagement, Hispanic infants were more likely to have a higher number of predicted visits and African American mothers a lower number when compared to White mothers. Practically, this difference may not be significant. The average number of visits for an African-American was 4.9 (*SD* = 4.9) while the mean for Caucasians was 5.4 (*SD* = 5.1) and for all non-African-American races the mean was 5.3 (*SD* = 5.2).

There are some Hispanic clients who do not speak English as a first language who receive the additional support of a community health educator (CHE). The role of the CHE is more of a peer mentor to help assist the caregiver with accessing and understanding services in combination with the medical care of the nurse. A follow-up analysis for this subgroup was conducted to examine the influence of the CHE on engagement. In this sample there were 20 Hispanic families who also had a CHE and there were 79 who did not have a CHE. These numbers were likely too small to detect a significant effect, but there appears to be a trend towards better engagement and maltreatment outcomes. About 95% of those with a CHE

received more than one home visit compared to 83% of non-CHE Hispanics (p = .15), a similar trend was found for number of total visits. For child maltreatment, 5% of Hispanic families with a CHE (1/20) had a later report compared to 15% of Hispanic families without a CHE. This was not a statistically significant (p = .22) association but could represent a clinically important service component for non-English speaking families.

Table 6

Multivariate Logistic Regression Model Predicting Initial Engagement

	Model 2: Risk Factors			Model 3: Service Variables			Model 4: Full Model		
	OR	OR 95% CI		OR	OR 95% CI		OR	OR 95% CI	
Infant Ethnicity (Hispanic=1)	1.58	0.93	2.69	**1.79**	**1.02**	**3.13**	1.56	0.91	2.65
Cumulative Risk Score				**1.11**	**1.07**	**1.15**			
Caregiver Risk									
History of Physical Abuse/Neglect	**2.41**	**1.26**	**4.63**				**2.53**	**1.31**	**4.89**
Cognitive/Learning Disabilities	**2.70**	**1.52**	**4.80**				**2.67**	**1.49**	**4.78**
Intended Pregnancy	**1.37**	**1.03**	**1.83**				**1.37**	**1.02**	**1.84**
Inadequate Prenatal Care	**1.29**	**1.03**	**1.62**				1.24	0.99	1.56
Chronic Physical Health Problem	**0.77**	**0.60**	**0.98**				**0.75**	**0.59**	**0.97**
Infant Risk									
Low Birthweight/Preterm	**1.41**	**1.12**	**1.77**				**1.42**	**1.12**	**1.79**
Prior CA/N Reported Victim	**0.50**	**0.29**	**0.86**				0.77	0.42	1.42
Age at Referral (Postnatal >90)									
Prenatal > 90 days birth				**2.05**	**1.40**	**2.98**	**1.77**	**1.16**	**2.68**
Prenatal < 90 days before birth				**5.27**	**3.32**	**8.38**	**4.58**	**2.81**	**7.47**
Postnatal <30 days after birth				**1.83**	**1.28**	**2.61**	**1.71**	**1.16**	**2.52**
Postnatal 30 - 90 days after birth				**1.60**	**1.00**	**2.56**	1.50	0.93	2.42
Referral Source (Self-Referral)									
Healthcare Agency				1.24	0.96	1.60	1.17	0.90	1.52
Social Service Agency				1.31	0.98	1.77	**1.39**	**1.03**	**1.88**
Referral to First Visit (>30 days)									
1 day				**2.72**	**1.74**	**4.25**	**2.48**	**1.60**	**3.85**
2-6 days				**1.92**	**1.40**	**2.65**	**1.86**	**1.35**	**2.56**
7-30 days				**1.53**	**1.13**	**2.06**	**1.49**	**1.11**	**2.02**

Note: Estimates in bold significant at $p < .05$.

Table 7

Multivariate Negative Binomial Regression Model Predicting Number of Home Visits (Retention)

	Model 2: Risk Factors			Model 3: Service Variables			Model 4: Full Model		
	Estimate	95% CI		Estimate	95% CI		Estimate	95% CI	
Infant Race (AA=1)	-0.06	-0.14	0.03	**-0.10**	**-0.19**	**-0.02**	**-0.10**	**-0.19**	**-0.01**
Infant Ethnicity (Hispanic=1)	**0.17**	**0.02**	**0.31**	**0.17**	**0.03**	**0.32**	**0.18**	**0.03**	**0.32**
Relationship to Infant (bio mother=1)	**-0.18**	**-0.34**	**-0.02**	**-0.18**	**-0.33**	**-0.03**	**-0.23**	**-0.38**	**-0.07**
Screeners									
Everyday Stress Index	**0.11**	**0.05**	**0.18**	**0.09**	**0.02**	**0.17**	**0.09**	**0.03**	**0.16**
Edinburgh Postpartum Depression	0.09	-0.02	0.19	**0.10**	**-0.01**	**0.21**	0.07	-0.04	0.18
Cumulative Risk Score				**0.06**	**0.05**	**0.07**			
Maternal Risk Factors									
Multiple Current Children	**-0.09**	**-0.15**	**-0.02**				-0.06	-0.12	0.00
Unemployed	-0.05	-0.11	0.00				**-0.07**	**-0.12**	**-0.01**
Current Domestic Violence	**0.32**	**0.21**	**0.44**				**0.31**	**0.20**	**0.42**
History of Physical Abuse/Neglect	**0.29**	**0.16**	**0.41**				**0.26**	**0.14**	**0.38**
Alcohol Use During Pregnancy	0.04	-0.09	0.17				0.01	-0.12	0.14
Mental Health Diagnosis	**0.13**	**0.06**	**0.20**				**0.13**	**0.06**	**0.20**
Cognitive/Learning Disabilities	**0.55**	**0.43**	**0.67**				**0.53**	**0.42**	**0.65**
Prior Fetal Death	**0.20**	**0.11**	**0.30**				**0.17**	**0.08**	**0.26**
Prior Pregnancy Complication	**0.30**	**0.16**	**0.44**				**0.28**	**0.15**	**0.42**
Inadequate Prenatal Care	**0.07**	**0.00**	**0.14**				**0.07**	**0.01**	**0.14**
<18 Months Between Pregnancies	**0.15**	**0.05**	**0.24**				0.08	-0.01	0.17
Chronic Physical Health Problem	**0.11**	**0.04**	**0.18**				**0.09**	**0.02**	**0.17**
Child Risk Factors									
Low Birthweight/Preterm	**0.14**	**0.08**	**0.20**				**0.18**	**0.11**	**0.24**
Drug-Exposed	**-0.14**	**-0.28**	**-0.01**				-0.09	-0.22	0.04
Heart/Lung Complications	0.11	0.00	0.21				**0.16**	**0.05**	**0.27**
Prior CA/N Reported Victim	**-0.31**	**-0.50**	**-0.12**				-0.10	-0.31	0.10
Infant Age at Referral (Postnatal >90)									

Prenatal > 90 days birth	**0.49**	**0.37**	**0.62**	**0.50**	**0.37**	**0.63**
Prenatal < 90 days before birth	**0.48**	**0.35**	**0.61**	**0.48**	**0.34**	**0.62**
Postnatal <30 days after birth	**0.15**	**0.03**	**0.27**	**0.21**	**0.08**	**0.34**
Postnatal 30 - 90 days after birth	**0.15**	**0.00**	**0.30**	**0.17**	**0.02**	**0.32**
Referral to First Visit (> 30 days)						
1 day	**0.48**	**0.36**	**0.61**	**0.41**	**0.28**	**0.53**
2-6 days	**0.32**	**0.21**	**0.42**	**0.27**	**0.17**	**0.37**
7-30 days	**0.22**	**0.12**	**0.31**	**0.18**	**0.09**	**0.27**

Note: Estimates in bold significant at $p < .05$.

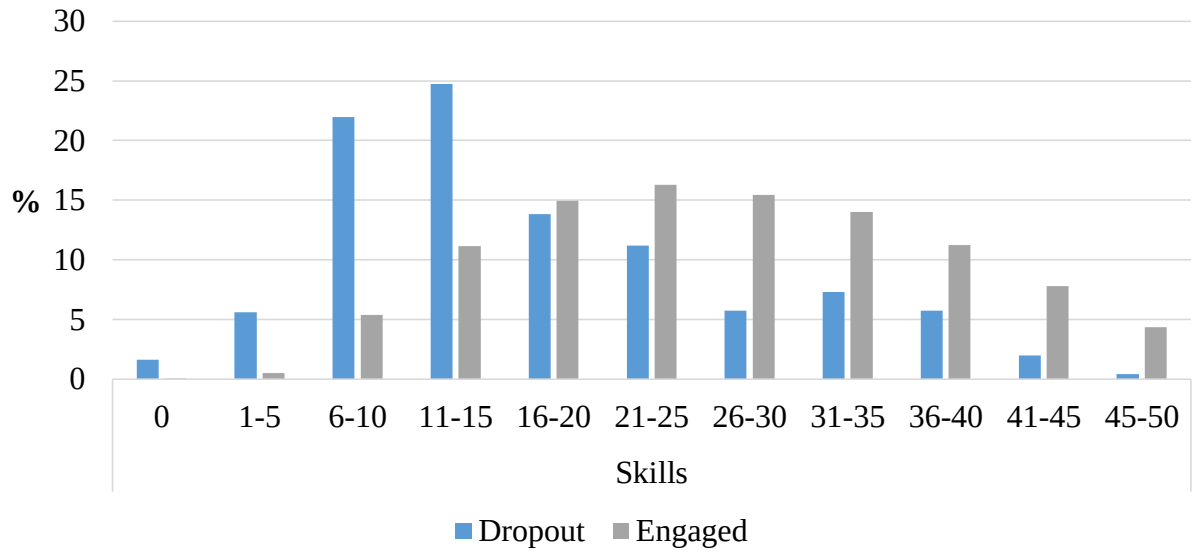
Q1.4: Is retention in services/dosage associated with documented skills training?

Nurses teach a number of parenting skills during their visits. The number of skills covered in a visit, however, will vary according to the capacity of the mother, length of the visit, and primary needs at the time of the visit. Thus this question explores whether there is a relationship between the number of visits (dosage) and the number of skills delivered by the nurse according to the case notes. This was first tested by examining the correlation between retention and the number of skills reported by the nurse. The correlation was positive and significant for number of visits ($r = .38$, $p < .001$), number of weeks of service contact ($r = .34$, $p < .001$), and total hours of nurse contact ($r = .39$, $p < .001$). Families who received more than one home visit had more documented skills ($M = 26.5$, $SD = 10.6$) compared to families who dropped out after the first visit ($M = 17.4$, $SD = 10.5$; $t = 21.09$, $p < .001$). Examining the distribution of skills in Figure 5 between these two initial engagement groups does show that while the difference is statistically significant, there are some families who drop out after the first visit that still receive a fairly high number of teaching skills in that first visit. Total skills was then regressed on total visits and a set of covariates. The results can be found in Table 8. The findings indicated that for each additional visit, a family can expect an average increase of .78 total skills ($\beta_{skills} = .77$, $p < .001$). Younger caregivers, those who initiate prenatally, first-time pregnancies, and those referred by a healthcare agency had higher predicted number of skills. There was also a significant negative association between the cumulative risk score and the number of skills received.

Table 8
Multivariate Regression Model Predicting Number of Skills

	Estimate	t	p
Intercept	25.18	18.55	<.0001
Number of Visits	0.77	27.41	<.0001
Caregiver Age	-0.16	-6.61	<.0001
First Visit Prenatal	2.32	6.26	<.0001
Prior Pregnancy	-0.82	-2.67	0.01
Race (AA)	-0.64	-1.71	0.09
Ethnicity (Hispanic)	0.07	0.08	0.93
Geographic Location			
Urban	-0.18	-0.29	0.77
Rural	-0.58	-0.78	0.43
Suburban	0.00	.	.
Caregiver Education			
GED/HS	-0.43	-1.14	0.26
College Grad	-0.04	-0.06	0.96
No HS	-0.40	-1.00	0.32
Some College	0.00	.	.
Referral Source			
Healthcare Referral	1.67	4.27	<.0001
Other Agency Referral	-0.76	-1.67	0.10
Self-Referral	0.00		
Cumulative Risk	-0.18	-2.71	0.01

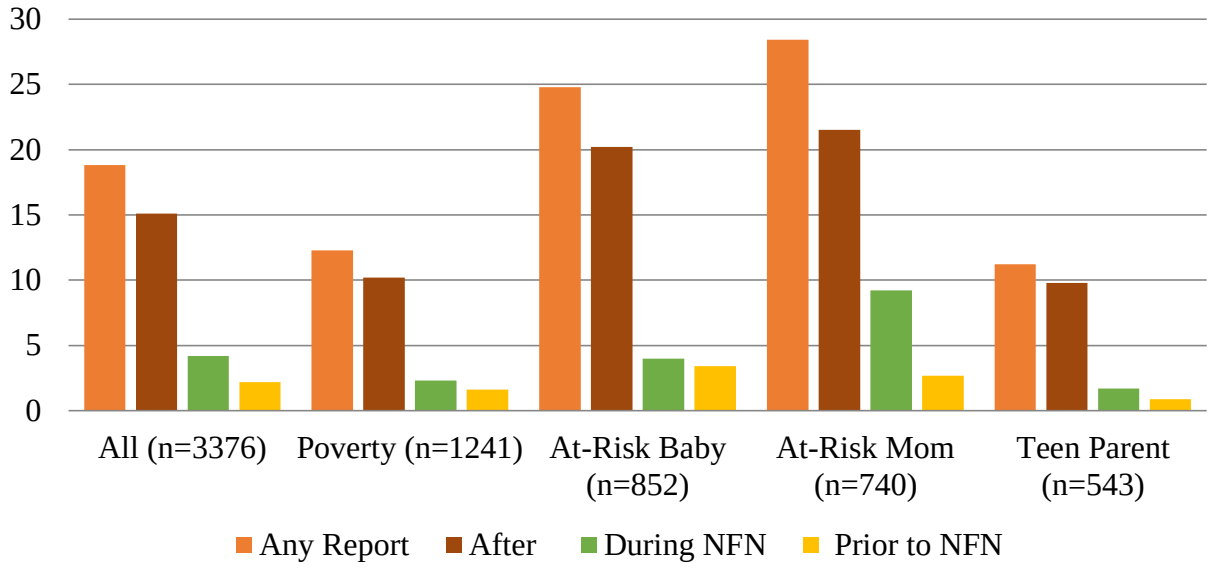
Figure 7: Distribution of Number of Skills Taught by the Nurse for Dropout and Engaged Group



Q1.5: Are services associated with risk for maltreatment? *If so, is there a dose response relationship?* The analysis seeks to determine to what extent a dose response exists between services and outcome. The assumption would be that families who receive more services will receive more of the protective effects of the intervention and thus a graded response between services and rates of report may exist.

The following Figure 8 describes the child abuse and neglect (CA/N) reporting history of families across service population. The percent of families with any reports ever is provided then the percentage by timing of a report before, during, or after NFN. Overall, 18.8% of the sample had a record of a CA/N report in the administrative database. However, not all of these reports occurred after the intervention. A little over two percent of families (2.2%) had a report prior to referral to NFN, 4.2% had a report between referral and termination, and 15.5% had a report following termination of services. There was a significant association between service population and report. The at-risk baby and at-risk mother subgroups had over twice the rate of reports after termination than the poverty and teen parent risk group.

Figure 8: Rate of CA/N Report by Report Timing and Service Subpopulation



Bivariate analysis for CA/N reports following termination. There were caregivers who enrolled prenatally that did not have a later postpartum visit and thus infant birth data was not collected (243 of 955 caregivers). All further analyses for CA/N outcomes are limited to only families who had at least one postpartum visit. This restriction excluded an additional 77 prenatally initiated caregivers from analysis. The final sample had 3,299 caregivers with 635 enrolled prenatally and 2,664 enrolled postpartum. This is sample group C in the client flow diagram.

Arguably NFN can do little to prevent CA/N prior to engagement and reports that occur during services often occur very early prior to any reasonable expectation of service impact. Further, reports that occur during services may have been initiated by the nurse after observing suspected abuse or neglect. This would potentially lead to some degree of surveillance bias for those families who remain in services longer. Therefore remaining analyses focus on reports following termination. Table 9 provides the basic descriptive analysis for the different service utilization measures and risk for a CA/N report during the follow-up period. There is also a column for a report with "qualifiers". This refers to reports that have the parent listed as the perpetrator, are not unsubstantiated, and are not sexual abuse type. These qualifiers represent maltreatment outcomes that are not necessarily the intended target of home visiting. Perpetrators who are not a parent, such as another family member, may have been out of the control of the primary caregiver receiving the services. Further, the target behaviors of home visiting are most related to child physical abuse and neglect. In some interventions that focus on improving parent-child interactions, treatment for sexual abuse prevention is actually contraindicated (California Evidence-Based Clearinghouse for Child Welfare, nd; Chadwick Center on Children and Families, 2004). Unsubstantiated cases were removed for this category as well as these may

represent inappropriate reports and low-risk reports where not even assessment-track services were indicated. In general, these qualifiers provide a more stringent threshold to be counted as an adverse outcome, but all associations are similarly significant and in the same direction. There were no differences in significance or effect for the predictors for the qualified CA/N outcome versus any report. All outcomes reported will focus on any report as opposed to the qualified CA/N outcome.

Table 9 compared the average number of prenatal, postpartum, and total visits and a report of CA/N. A t test was used to compare the root-transformed number of visits for families with a CA/N report compared to those without due to the non-normal count distribution of the visit predictor. The results indicate that families with a CA/N report received on average fewer prenatal visits, more postpartum visits, and a similar number of total visits. Table 9 provides further breakdown by visit categories and rate of CA/N report. For prenatal visits, 18.8% of those who had no prenatal visits (only postpartum visits) had a later CA/N report compared to a report rate of 1.1% for families that received any number of prenatal visits. Only eight of the 510 families with a later CA/N report (and had a valid child birth date to match with the state administrative system) received a prenatal visit.

The relationship between service dose and reports is very different when looking at postpartum visits. Among families receiving postpartum visits, the rate of CA/N increases from 13.0% for those with one visit to 15.0% for those with 2-3 visits, and 16.7% for those with four or more visits. A post hoc trend test was conducted using the Cochran-Armitage Trend test and found a significant effect ($Z = -2.29$, $p < .05$) for the raw number of postpartum visits and frequency of later CA/N report. If the families who engage prenatally are included there is no

relationship between visits and later report. But for postpartum-only families, a higher number of visits is associated with greater risk of report.

Table 9
Association between Service Utilization Characteristics and Rate of Later CA/N Report

	No CA/N Report (n=2,789)	Follow-Up CA/N (n=510)	Follow-Up CA/N with Qualifiers (n=323)	sig
Sample Percentage	84.5	15.5	9.6	
Infant Age at Referral				*
Prenatal > 90 days birth	94.7	5.4	3.5	
Prenatal < 90 days before birth	96.2	3.8	2.3	
Postnatal <30 days after birth	83.2	16.8	10.7	
Postnatal 30 - 90 days after birth	71.9	28.0	17.8	
Postnatal > 90 days after birth	69.6	30.4	19.0	
Referral Source				*
Self-Referral	88.6	11.4	5.7	
Healthcare Agency	85.2	14.8	9.3	
Social Service Agency	79.7	20.3	14.1	
Time from Referral to First Visit				*
1 day	80.6	19.4	12.0	
2-6 days	84.7	15.3	9.7	
7-30 days	85.9	14.0	9.0	
>30 days	81.6	18.4	11.4	
Prenatal Visits	0.6 (1.5)	0.1 (0.5)	0.1 (0.6)	*
Postpartum Visits	4.8 (4.7)	5.3 (4.9)	5.3 (5.0)	*
Total Nurse Visits	5.5 (5.2)	5.3 (5.0)	5.4 (5.1)	ns
Prenatal Nurse Visits Categorical				*
0 Visits	81.2	18.8	11.9	
1 Visit	98.5	1.5	0.5	
2-3 Visits	98.7	1.3	1.3	
4+ Visits	99.0	1.0	1.0	
Postpartum Nurse Visits Categorical				*
1 Visit	87.0	13.0	8.4	
2-3 Visits	85.0	15.0	9.4	
4+ Visits	83.3	16.7	10.6	
Total Nurse Visits Categorical				ns
1 Visit	85.3	14.7	9.4	
2 Visits	82.5	17.6	10.4	
3-4 Visits	83.6	16.4	10.5	
5-6 Visits	86.4	13.6	8.4	
7+ Visits	84.9	15.1	9.8	
Total Hours with Nurse	7.2 (6.7)	7.1 (6.5)	7.2 (6.7)	ns
Total Days from First to Last Visit	110.1 (145.3)	105.0	105.0 (143.7)	ns

		(144.7)		
Skills Covered by Nurse	26.3 (10.6)	23.6 (10.6)	23.4 (10.6)	*
Termination Reason				*
Refused Services	84.5	15.6	9.9	
Moved/Cannot Locate	81.0	18.9	15.3	
No Need/Duplicate Services	89.5	10.5	4.7	
Completed 24 months	70.3	29.7	13.5	
Other Reason	85.1	14.9	9.2	

Note: *$p < .05$

There are several other associations between service utilization and CA/N report from the bivariate table. There was a significant association between the timing of the referral and later report. A Cochran-Armitage Trend test detected a significant trend in that the later the referral came relative to the birth of the child, the more likely a family was to have a later report ($Z = -11.9$, $p < .0001$). While 4.9% of families with a referral for NFN services 90 days before birth had a later report 30.4% of those with a referral coming 90 days after the infant was born had a later report. Families referred from a social service agency were more likely to have reports compared to healthcare providers or self-referrals. Families with a later report were more likely to have been terminated because they moved or the nurse could not locate them and received on average less skills compared to those without a later report.

Multivariate model for CA/N report. The outcome of CA/N report was regressed on blocks of predictor covariates in a manner similar to what was reported for engagement and retention. First, child and caregiver demographic factors were entered, followed by the set of unique risk factors, and then service variables. A Cox regression model (Table 10, n = 3,299) was used including clustering at the nurse level estimating the time-to-event for a first CA/N report. Results of the full model will be discussed.

Table 10
Multivariate Cox Regression Model Predicting CA/N Report

	Model 1: MCH and Scales			Model 2: Risk Factors			Model 3: Service Variables			Model 4: Full Model		
	HR	HR 95% CI		HR	HR 95% CI		HR	HR 95% CI		HR	HR 95% CI	
Infant Race (AA=1)	**0.55**	**0.45**	**0.66**	**0.56**	**0.46**	**0.68**	**0.62**	**0.51**	**0.75**	**0.64**	**0.52**	**0.78**
Infant Ethnicity (Hispanic=1)	0.85	0.50	1.45	0.98	0.56	1.71	0.88	0.51	1.50	0.93	0.54	1.58
Infant Gender (Female=1)	0.95	0.80	1.12	0.96	0.81	1.14	0.92	0.78	1.08	0.94	0.79	1.11
Caregiver Relationship to Infant (bio mother=1)	1.06	0.73	1.55	1.34	0.85	2.12	**1.64**	**1.09**	**2.47**	**1.92**	**1.23**	**2.99**
Caregiver Marital Status (Married=1)	0.83	0.66	1.04	**0.80**	**0.65**	**0.99**	**0.75**	**0.61**	**0.94**	**0.81**	**0.66**	**1.00**
Zip level Urbanicity												
(Rural)												
Urban	0.84	0.65	1.08	0.91	0.75	1.10	1.03	0.77	1.39	1.03	0.78	1.35
Suburban	0.75	0.53	1.05	0.89	0.66	1.20	1.02	0.70	1.47	1.03	0.71	1.48
Screeners												
Ages and Stages	0.99	0.79	1.23	0.92	0.76	1.11	0.93	0.73	1.18	0.96	0.75	1.21
Everyday Stress Index	**1.58**	**1.23**	**2.03**	**1.54**	**1.28**	**1.86**	**1.59**	**1.22**	**2.07**	**1.52**	**1.19**	**1.93**
Edinburgh Postpartum Depression	0.76	0.50	1.16	0.85	0.56	1.29	0.73	0.48	1.10	0.71	0.46	1.11
Cumulative Risk Score	**1.16**	**1.12**	**1.19**	-	-	-	**1.14**	**1.10**	**1.18**	-	-	-
Maternal Risk Factors												
Homeless				0.94	0.64	1.37				1.15	0.72	1.83
Father Unknown/Not Involved				0.91	0.51	1.64				0.98	0.57	1.68
Prior DFS Involvement				**1.39**	**1.02**	**1.90**				1.25	0.92	1.71
Teenage Mother				1.02	0.83	1.26				1.12	0.90	1.39
No High School Education				**1.43**	**1.18**	**1.73**				**1.44**	**1.17**	**1.76**
Multiple Current Children				**1.53**	**1.23**	**1.90**				**1.42**	**1.15**	**1.77**
Unemployed				**1.67**	**1.36**	**2.03**				**1.63**	**1.33**	**2.00**
Current Domestic Violence				0.83	0.52	1.32				0.80	0.49	1.29

History of Rape/Sexual Abuse	1.09	0.67	1.79				1.03	0.61	1.76
History of Physical Abuse/Neglect	1.10	0.74	1.64				1.25	0.83	1.87
Alcohol Use During Pregnancy	0.54	0.28	1.06				**0.52**	**0.28**	**0.98**
Drug Use During Pregnancy	1.16	0.82	1.64				1.17	0.83	1.65
Smoking During Pregnancy/In Home	1.04	0.84	1.30				1.04	0.85	1.28
Mental Health Diagnosis	**1.33**	**1.04**	**1.70**				**1.40**	**1.09**	**1.79**
Cognitive/Learning Disabilities	**1.72**	**1.06**	**2.77**				1.49	0.91	2.45
Neurological Impairments/Injury	0.88	0.46	1.67				0.93	0.48	1.80
Intended Pregnancy	1.01	0.79	1.31				0.94	0.73	1.22
Prior Low Birthweight/Preterm	**0.31**	**0.15**	**0.67**				**0.30**	**0.14**	**0.66**
Prior Fetal Death	**0.59**	**0.42**	**0.85**				**0.64**	**0.45**	**0.92**
Prior Pregnancy Complication	0.81	0.43	1.51				0.89	0.45	1.76
Current Pregnancy Complication	1.06	0.77	1.46				1.13	0.81	1.56
Inadequate Prenatal Care	1.10	0.90	1.34				1.10	0.90	1.36
<18 Months Between Pregnancies	0.80	0.57	1.14				0.93	0.66	1.31
Chronic Physical Health Problem	0.99	0.76	1.28				1.03	0.78	1.35
Child Risk Factors									
Special Care/NICU	1.16	0.75	1.78				1.03	0.68	1.55
Low Birthweight/Preterm	**1.30**	**1.08**	**1.57**				1.11	0.92	1.34
Drug-Exposed	**2.04**	**1.47**	**2.83**				**1.97**	**1.40**	**2.78**
Jaundice	**1.56**	**1.03**	**2.36**				1.53	0.96	2.43
Heart/Lung Complications	0.80	0.55	1.16				0.69	0.47	1.01
Major Congenital Disability	**1.99**	**1.21**	**3.26**				**1.78**	**1.07**	**2.94**
Prior CA/N Reported Victim	**2.26**	**1.47**	**3.50**				1.19	0.75	1.90
Infant Age at Referral									
(Prenatal > 90 days birth)				1.00			1.00		
Prenatal < 90 days before birth				0.78	0.39	1.59	0.77	0.38	1.56
Postnatal <30 days after birth				**5.49**	**3.42**	**8.81**	**5.03**	**3.20**	**7.92**
Postnatal 30 - 90 days after birth				**7.87**	**4.93**	**12.55**	**7.87**	**4.78**	**12.97**

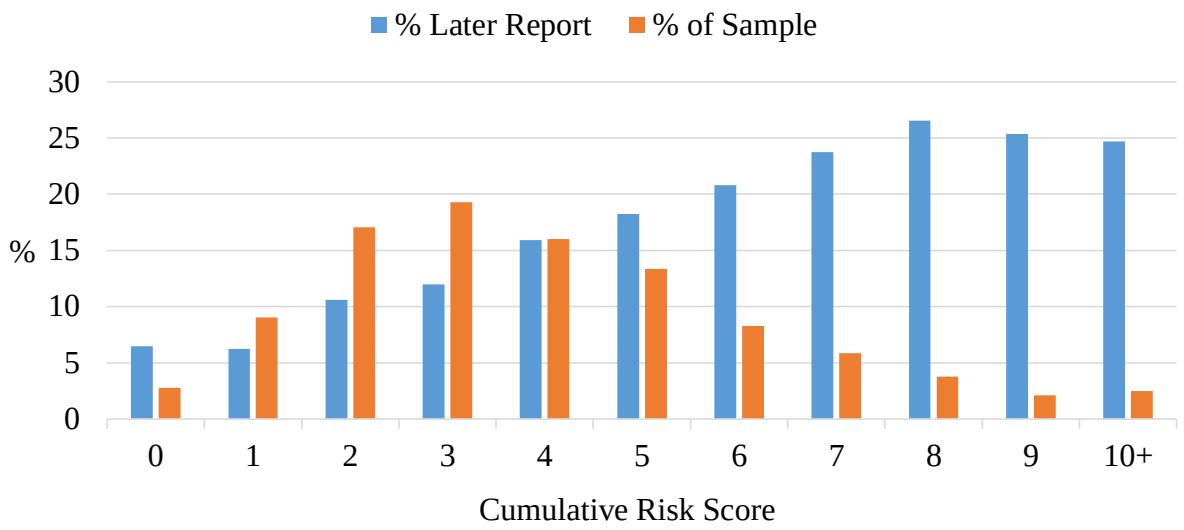
Postnatal > 90 days after birth	**8.14**	**4.61**	**14.37**	**8.52**	**4.74**	**15.32**
Referral Source						
(Self-Referral)	1.00			1.00		
Healthcare Agency	0.87	0.68	1.12	0.99	0.78	1.27
Social Service Agency	1.15	0.84	1.59	1.11	0.79	1.56
Time from Referral to First Visit						
(1 day)	1.00			1.00		
2-6 days	0.80	0.57	1.12	0.83	0.59	1.15
7-30 days	0.88	0.64	1.22	0.92	0.67	1.26
>30 days	**2.16**	**1.47**	**3.18**	**2.45**	**1.73**	**3.48**
Number of Total Visits (root-transformed)	1.10	0.98	1.25	**1.22**	**1.07**	**1.40**

Note: HR in bold are significant at $p < .05$

Of interest in this section primarily is the relationship between service dosage and later CA/N report while controlling for other covariates in the model. When controlling for the level of cumulative risk, total number of visits was not associated with report, but when individual risk factors were included in the model, the total number of visits (largely postpartum visits) was associated with an increased risk for later report. African-American children had a lower risk for later report compared to children of other races in the sample. Caregiver stress as measured by the ESI was associated with an increased risk for CA/N. A caregiver above the mean ESI level in this sample was 52% more likely to have a later CA/N report. No high school education, being unemployed, having other children in the home, and having a mental health diagnosis was associated with increased risk among maternal risk factors. Drug-exposed infants and those with major disabilities had 97% and 78% greater risk for CA/N. Families referred later in the postpartum period and those with a longer delay from referral to first visit had a higher risk for CA/N compared to those enrolled prenatally with a same day referral and first visit.

Multivariate models indicate that each increase in the cumulative risk score increased the risk for CA/N 10-18%. The follow figure displays the bivariate relationship between risk and report by displaying the rate of later CA/N report from zero to ten or greater on the cumulative risk scale based on the 37 risk factors identified for this study. The relationship between risk score and rate of report has a strong linear relationship ($R^2 = .94$).

Figure 9. Relationship between Cumulative Risk Score and Rate of Later Maltreatment Report



The next piece of the multivariate analysis was to consider level of service dosage as a moderator of other covariates in predicting risk for CA/N report. Several a priori groups were analyzed. First, it was hypothesized that there would be a significant interaction between level of cumulative risk and service level in that families with higher risk who received more visits would have a decreased risk compared to high risk families who received a lower doses of visits. After descriptive analyses indicated a different direction of effect for prenatal and postpartum visits, these were analyzed separately.

A dosage by level of risk interaction term was entered in the final model and was not significant when entered as a continuous measure of risk or in categorical levels of risk. When examining service population, there were no significant interaction terms examining number of visits as a moderator. Next, CA/N report was regressed on child and caregiver characteristics and individual risk factors with interaction terms for visits. There were no significant interaction effects by dosage for child and caregiver demographics but there were two effects for specific risk factors.

First, there was a significant interaction effect for the maternal risk factor of having a prior low birth weight baby ($\beta_{visits\ x\ priorlbw} = -0.27$, $p < .01$). This finding suggests that as the number of visits increases, the risk for a later CA/N report among mothers who had a previous low birthweight baby compared to those without a prior is smaller, after controlling for a variety of other factors. The second significant interaction effect was for infant who had a prior CA/N report prior to NFN services ($\beta_{visits\ x\ priorcan} = -1.56$, $p < .01$). Again, the relationship between number of visits and later CA/N report was different depending on whether the infant had a prior CA/N report. Interaction terms are best interpreted graphically. The following figures (Figure 10 and Figure 11) provide the raw bivariate rates of later CA/N report based on three levels of postpartum visits. For the general NFN population, an increase in number of visits is associated with a higher risk for later CA/N report. However, for both prior low birthweight caregivers and prior CA/N report infants, a higher dosage of treatments appears to have a protective effect for later reports in multivariate models.

At the bivariate level, prior CA/N report was associated with an increased risk for later report as 35% of these families did indeed have another recurrent report. However, there was a significant interaction between number of visits and risk. The rate of CA/N decreased for those with 1 visit (52%), 2-3 visits, (30%), and 4 or more visits (26%). This was the only truly linear dosage effect with visits for any group in the current study. The more visits that a family received, the less likely they were to have a future report.

Figure 10. Rate of Later CA/N Report by Prior Low Birthweight Baby and Number of Visits

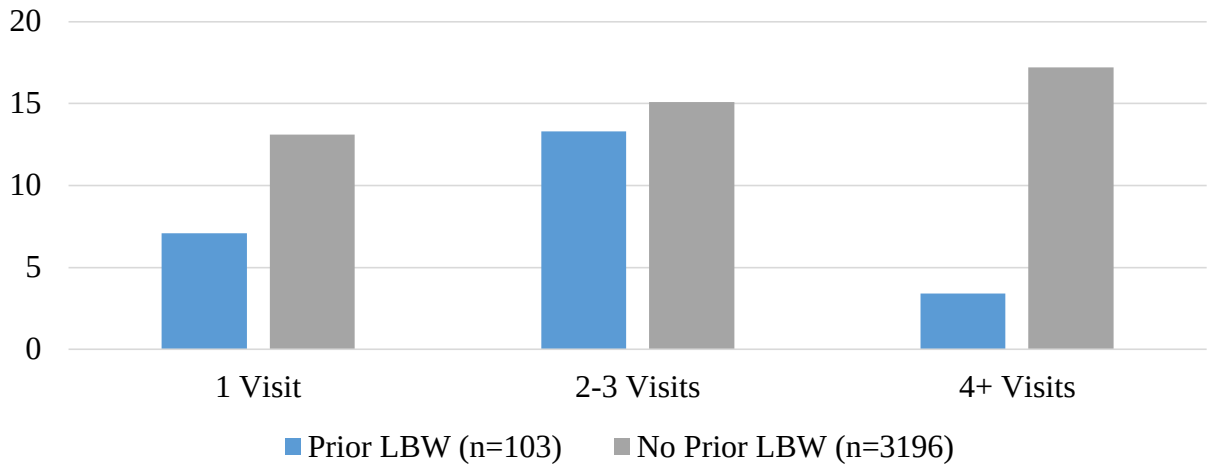
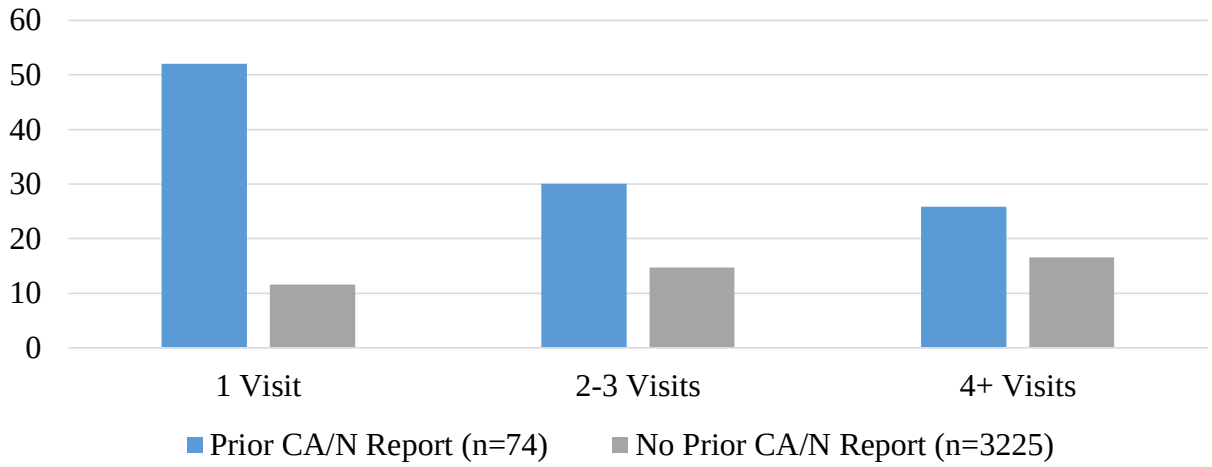


Figure 11. Rate of Later CA/N Report by Prior Infant CA/N Report and Number of Visits



Prenatal versus postpartum service initiation. Given findings indicating a sharp difference in CA/N outcomes for families who receive their first visit during the prenatal period compared to those who begin services after the infant was born, further analyses were conducted for these groups separately. Prenatal versus postpartum groups are explored further in the subgroups section describing their outcomes for ASQ, ESI, EPDS, and CA/N reports. Note that

this bivariate analysis refers back to the full sample (n=3,620) and is not limited to those with birth information or at least one postpartum visit (sample A).

Table 11 provides a comparison of the child and family characteristics for prenatal versus postpartum families. Prenatal families, compared to those with only postpartum visits, were more likely to be African-American (63.6% vs. 41.3%, $\chi^2 = 108.7$, $p < .001$), the primary caregiver was more likely to be the biological mother (98.9% vs. 96.0%, $\chi^2 = 14.5$, $p < .001$), less likely to have a partner (11.9% vs. 19.4%, $\chi^2 = 21.4$, $p < .001$), more likely to be a teenage mother (43.4% vs. 34.1%, $\chi^2 = 21.0$, $p < .001$), the infant was less likely to be born low birth weight (10.2% vs. 19.1%, $\chi^2 = 30.3$, $p < .001$) or premature (13.5% vs. 19.9%, $\chi^2 = 14.5$, $p < .001$).

Table 11
Demographic Characteristics of Families Initiating Visits Postpartum versus Prenatal

	Total Sample (n=3,620)	Postpartum Initiated (n=2665)	Prenatal Initiated (n=955)	sig
Infant Characteristics	100.0	73.6	26.4	
Child Age at Follow-up (years)	2.9 (.8)	2.9 (.8)	2.8 (.6)	*
Infant Race				*
Black	46.3	41.3	64.1	
White	46.2	50.6	30.6	
Biracial	6.7	7.1	5.0	
Infant Ethnicity (Hispanic)	3.7	3.7	4.0	ns
Infant Gender (female)	48.4	49.3	45.0	*
Low Birthweight (<2500 grams)	17.2	19.1	10.2	*
Very Low Birthweight (<1500 grams)	2.5	2.9	1.2	*
Preterm Birth (<37 weeks gestation)	18.6	19.9	13.6	*
Caregiver Characteristics				
Relationship to Infant (bio mother)	96.9	96.0	99.2	*
Age at Referral (<20 years)	36.1	34.1	41.5	*
Zip level Urbanicity				ns
Urban	78.1	78.0	78.3	
Rural	13.8	13.7	14.3	
Suburban	8.1	8.3	7.5	
Previous Pregnancy	52.7	51.7	55.3	ns
Living Children	41.3	43.4	35.6	*
Marital Status				*
Single	78.9	76.6	85.3	
Married/Consensual Union	17.4	19.4	11.6	
Separated	1.7	2.0	1.0	
Divorced	1.7	1.6	1.9	
Level of Education				*
No HS	37.3	35.4	42.5	
HS/GED	37.3	38.3	34.5	
Some College	19.9	19.9	19.9	
College Degree	5.6	6.5	3.1	
Employment Status				*
Unemployed	50.6	49.5	53.5	
Disabled	2.0	2.1	1.8	
Homemaker	8.0	9.5	3.8	
Student	15.9	14.1	20.8	
Part-Time	10.8	10.9	10.4	
Full-Time	12.5	13.7	9.4	

Note: *$p < .05$

Table 12 compares risk factors and Table 13 compares service utilizations patterns for prenatal versus postpartum initiators. Referral source was strongly associated with when a caregiver initiated services ($\chi^2 = 1024.9$, $p < .001$) with over half of postpartum visits coming from the hospital setting and almost half of prenatal referrals from self or family referral. The prenatal mothers were more likely to be homeless (10.1% vs. 3.5%, $\chi^2 = 53.9$, $p < .001$), currently experiencing partner violence (7.7% vs. 4.3%, $\chi^2 = 14.0$, $p < .001$), have a personal history of child abuse or neglect (8.6% vs. 3.3%, $\chi^2 = 36.9$, $p < .001$), have a prior fetal death (13.5% vs. 6.9%, $\chi^2 = 31.5$, $p < .001$), and have less than 18 months between their last pregnancy (16.4% vs. 7.3%, $\chi^2 = 56.3$, $p < .001$). Those who initiate services postpartum are more likely have a child in the NICU or special care nursery (4.8% vs. 2.7%, $\chi^2 = 5.9$, $p < .01$), are more likely to have a small or early baby (27.3% vs. 18.4%, $\chi^2 = 23.4$, $p < .001$), and more likely to have a drug exposed baby (9.8% vs. 1.8%, $\chi^2 = 48.2$, $p < .001$).

Table 12
Caregiver and Child Risk Factors for NFN Sample Comparing Families who Initiate Visits Postpartum versus Prenatal

	Total Sample (n=3,620)	Postpartum Initiated (n=2665)	Prenatal Initiated (n=955)	
Caregiver Risk Factors	100.0	38.4	23.5	sig.
Psychosocial				
Homeless	5.4	3.5	10.8	*
Father Unknown/Not Involved	2.5	2.3	3.0	ns
Prior DFS Involvement	7.1	7.8	5.1	*
Teenage Mother	36.1	34.1	41.5	*
No High School Education	36.6	34.6	42.2	*
Multiple Current Children	41.3	43.4	35.6	*
Unemployed	49.9	48.7	53.3	*
Violence Exposure				
Current Domestic Violence	5.2	4.3	7.6	*
History of Rape/Sexual Abuse	3.6	3.0	5.2	*
History of Physical Abuse/Neglect	4.4	3.3	7.5	*
Behavioral Health				
Alcohol Use During Pregnancy	4.6	4.4	5.1	ns

Drug Use During Pregnancy	13.3	14.8	9.0	*
Smoking During Pregnancy/In Home	24.4	24.3	24.8	ns
Mental Health Diagnosis	20.0	20.0	19.8	ns
Cognitive/Learning Disabilities	4.6	4.5	4.7	
Neurological Impairments/Injury	2.0	2.0	1.9	ns
Maternal Health				
Unintended Pregnancy	85.4	84.3	88.6	*
Prior Low Birthweight/Preterm	3.2	2.9	4.3	*
Prior Fetal Death	8.4	6.9	12.4	*
Prior Pregnancy Complication	3.2	2.6	5.0	*
Current Pregnancy Complication	10.2	10.5	9.2	ns
Inadequate Prenatal Care	21.6	21.6	21.6	ns
<18 Months Between Pregnancies	9.4	7.3	15.2	*
Chronic Physical Health Problem	14.5	14.2	15.3	ns
Child Risk Factors				
Special Care/NICU	4.0	4.8	2.0	*
Low Birthweight/Preterm	23.7	27.3	13.8	*
Drug-Exposed	7.6	9.8	1.4	*
Jaundice	2.8	3.6	0.3	*
Heart/Lung Complications	6.1	7.7	1.9	*
Major Congenital Disability	2.2	2.7	0.9	*
Prior CA/N Reported Victim	2.0	2.8	0.0	*
Cumulative Risk Score	4.6 (2.3)	4.6 (2.4)	4.7 (2.3)	ns
Child Cumulative Risk Score	0.5 (0.7)	0.6 (0.8)	0.2 (0.5)	*
Caregiver Cumulative Risk Score	4.2 (2.1)	4.1 (2.1)	4.5 (2.1)	*

Note: *$p < .05$

Table 13
Service Characteristics Comparing Families Initiating Visits Postpartum Versus Prenatal

	Total Sample (n=3,620)	Postpartum Initiated (n=2665)	Prenatal Initiated (n=955)	sig.
Infant Age at Referral				*
Prenatal > 90 days birth	19.8	3.6	64.9	
Prenatal < 90 days before birth	11.9	3.6	35.0	
Postnatal <30 days after birth	55.9	75.9	0.1	
Postnatal 30 - 90 days after birth	5.9	8.0	0.0	
Postnatal > 90 days after birth	6.6	8.9	0.0	
Referral Source				*
Self-Referral	18.3	10.8	39.5	
Healthcare Agency	57.5	70.0	22.6	
Social Service Agency	23.8	18.8	37.8	
Time from Referral to First Visit				*
1 day	9.9	9.0	12.3	
2-6 days	37.2	43.0	20.8	
7-30 days	43.2	38.7	55.7	
>30 days	9.8	9.3	11.2	
Prenatal Nurse Visits				*
0 Visits	73.6	100.0	0.0	
1 Visit	9.7	0.0	36.9	
2-3 Visits	9.8	0.0	37.0	
4+ Visits	6.9	0.0	26.2	
Postpartum Nurse Visits				*
0 Visits	8.8	0.0	33.4	
1 Visit	18.8	22.3	8.9	
2-3 Visits	27.2	30.6	17.6	
4+ Visits	45.3	47.1	40.1	
Total Nurse Visits				*
1 Visit	20.7	22.3	16.3	
2 Visits	15.8	17.3	11.7	
3-4 Visits	23.8	25.5	19.3	
5-6 Visits	14.3	14.0	15.1	
7+ Visits	25.4	21.0	37.6	
Termination Reason				*
Refused Services	31.7	32.5	29.7	
Moved/Cannot Locate	7.0	6.1	9.4	
No Need/Duplicate Services	2.5	2.7	1.9	
Completed 24 months	1.0	1.0	1.3	
Other Reason	57.7	57.7	57.7	

Note: *$p < .05$

Table 14 provides a comparison of models predicting service retention (number of visits) for postpartum versus prenatal families. Prenatal mothers receive more visits if there is not a father involved and the pregnancy was not intended. There is a stronger positive effect for prenatal mothers for children born low birthweight and those with heart or lung complications. This likely indicates that once the child is born, those with complications receive additional visits. Increased stress, health factors related to the pregnancy, and shorter delay between referral and the visit were significant factors in predicting increased visits for postpartum caregivers but not for prenatal. Both groups received more home visits if there was current domestic violence or a history of maternal child abuse or neglect. Postpartum mothers with mental health problems had higher numbers of visits but this was not significant for prenatal initiators.

Table 14
Multivariate Models Predicting Number of Visits Comparing Families Initiating Services during Postpartum versus Prenatal Period

	Postpartum Initiated (n=2665)			Prenatal Initiated (n=955)		
	Estimate	95% CI		Estimate	95% CI	
Infant Race (AA=1)	**-0.10**	**-0.18**	**-0.02**	-0.09	-0.27	0.08
Infant Ethnicity (Hispanic=1)	**0.26**	**0.10**	**0.41**	0.06	-0.24	0.35
Zip level Urbanicity						
Urban	0.00	-0.13	0.13	0.08	-0.20	0.35
Rural	**0.15**	**0.00**	**0.30**	0.12	-0.18	0.41
(Suburban)	0.00			0.00		
Screeners						
Ages and Stages	-0.04	-0.30	0.21	-0.05	-0.22	0.12
Everyday Stress Index	**0.12**	**0.04**	**0.20**	0.05	-0.08	0.17
Edinburgh Postpartum Depression	0.13	-0.02	0.28	0.01	-0.21	0.22
Cumulative Risk Score[1]	**0.06**	**0.01**	**0.04**	**0.09**	**0.01**	**0.06**
Maternal Risk Factors						
Father Unknown/Not Involved	0.04	-0.15	0.23	**0.38**	**0.09**	**0.67**
Multiple Current Children	**-0.12**	**-0.19**	**-0.05**	0.12	-0.01	0.25
Unemployed	-0.05	-0.11	0.02	**-0.15**	**-0.25**	**-0.04**
Current Domestic Violence	**0.34**	**0.20**	**0.48**	**0.25**	**0.05**	**0.45**
History of Rape/Sexual Abuse	0.06	-0.11	0.23	-0.04	-0.27	0.20
History of Physical Abuse/Neglect	**0.24**	**0.08**	**0.40**	**0.26**	**0.07**	**0.45**
Mental Health Diagnosis	**0.14**	**0.06**	**0.21**	0.13	-0.01	0.26
Cognitive/Learning Disabilities	**0.55**	**0.42**	**0.68**	**0.47**	**0.24**	**0.71**
Neurological Impairments/Injury	0.16	-0.04	0.37	0.08	-0.29	0.45
Unintended Pregnancy	0.03	-0.06	0.11	**0.23**	**0.06**	**0.40**
Prior Fetal Death	**0.20**	**0.09**	**0.32**	0.15	-0.01	0.31
Prior Pregnancy Complication	**0.37**	**0.19**	**0.54**	0.16	-0.07	0.39
Inadequate Prenatal Care	**0.12**	**0.05**	**0.20**	-0.02	-0.15	0.10
Chronic Physical Health Problem	**0.10**	**0.02**	**0.19**	0.03	-0.11	0.17
Child Risk Factors						
Low Birthweight/Preterm	**0.15**	**0.07**	**0.22**	**0.31**	**0.16**	**0.45**
Heart/Lung Complications	0.16	0.05	0.27	**0.39**	**0.03**	**0.76**
Time from Referral to First Visit						
1 day	**0.36**	**0.22**	**0.50**	0.20	-0.03	0.43
2-6 days	**0.18**	**0.07**	**0.29**	0.14	-0.05	0.33
7-30 days	0.06	-0.05	0.17	0.11	-0.06	0.27
>30 days	0.00			0.00		

Note: Estimates in bold significant at $p < .05$.
[1]Cumulative risk score estimate is from model without unique risk factors

In comparing predictors of later CA/N reports, the prenatal sample was limited to those with birth information (n=712, sample B). In this analysis, some of the prenatal families only have prenatal visits and some have postpartum visits later. Table 15 presents the results of separate models predicting later CA/N reports for prenatal versus postpartum families. Since the event of a CA/N report is rare (~1%) and the sample size is relatively small for the prenatal visit group, standard maximum likelihood estimation techniques likely yield biased results. Firth's penalized likelihood method was used to reduce the small sample bias and produce consistent estimates given the issues of quasi-complete separation that exist due to the rare event of CA/N reports in this subsample. Despite the corrections to the model specification, there standard errors and confidence intervals of the hazard ratios are large and significant effects were difficult to detect. Again, only 1% of the 712 prenatally engaged families had a later report of CA/N and significant inferential trends among these eight families likely do not exist. The only significant effect for the prenatal group was for number of postpartum visits that was received. Among the eight families with a later report, one client had one visit, and the rest had over four visits (max of 20). This is likely due to the fact that those who dropped out after one visit were more likely to not later provide infant birth information making the administrative data match impossible.

Table 15
Multivariate Model Predicting CA/N Report Comparing Families Initiating Services during Postpartum versus Prenatal Period

	Postpartum Initiated (n=2665)			Prenatal Initiated (n=712)		
	HR	95% CI Lower		HR	95% Profile Likelihood	
Infant Race (AA=1)	**0.64**	**0.52**	**0.78**	2.62	0.25	28.74
Infant Ethnicity (Hispanic=1)	0.82	0.48	1.40	10.65	0.94	76.39
Infant Gender (Female=1)	0.93	0.78	1.12	1.02	0.23	4.05
Caregiver Marital Status (Married=1)	**0.79**	**0.64**	**0.97**	1.03	0.10	5.86
Zip level Urbanicity						
(Rural)	1.00					
Urban	0.94	0.71	1.24	0.09	0.01	1.07
Suburban	0.93	0.65	1.33	0.84	0.07	6.71
Screeners						
Ages and Stages	0.94	0.53	1.64	0.47	0.05	2.41
Everyday Stress Index	**1.45**	**1.19**	**1.78**	0.82	0.17	3.83
Edinburgh Postpartum Depression	0.84	0.52	1.36	0.28	0.00	2.95
Cumulative Risk Score[1]	**1.14**	**1.10**	**1.17**	1.18	0.87	1.61
Maternal Risk Factors						
No High School Education	**1.43**	**1.19**	**1.72**			
Multiple Current Children	**1.49**	**1.20**	**1.85**			
Unemployed	**1.63**	**1.34**	**2.00**			
Prior Low Birthweight/Preterm	**0.31**	**0.14**	**0.68**			
Prior Fetal Death	**0.66**	**0.46**	**0.95**			
Child Risk Factors						
Drug-Exposed	**1.81**	**1.30**	**2.53**			
Time from Referral to First Visit						
(1 day)						
2-6 days	0.79	0.56	1.10	2.49	0.33	29.99
7-30 days	0.89	0.65	1.24	1.65	0.22	19.87
>30 days	1.23	0.88	1.72	0.94	0.01	23.30
Number of Postpartum Visits	**1.25**	**1.09**	**1.43**	**2.40**	**1.16**	**5.26**
Number of Prenatal Visits	-	-	-	0.64	0.12	2.92

Note: Hazard ratios in bold significant at $p < .05$
[1]Cumulative risk score estimate is from model without unique risk factors

Mediation analysis. In the main effect models represented, there appears to be a strong main effect of the number of visits on CA/N report. Further analysis was conducted to examine the effect of a third variable, cumulative risk, on this relationship. The simplest way to test this relationship is to explore the change in the effect for number of visits before and after controlling

for risk. In a model with CA/N regressed only on root-transformed number of postpartum visits, the HR = 1.22 (95% CI 1.08 to 1.38). If the cumulative risk scale variable is added to the model, the effect for postpartum visits is no longer significant (HR = 1.11, 95% CI 0.97 to 1.26). This indicates that the relationship between visits and CA/N report is no longer significant when controlling for the overall level of risk for a family.

A mediation model was also tested using the PROCESS macro (Hayes, 2012), selected for its ability to test mediation of dichotomous outcomes. The purpose of this analysis was to determine the direct effect of cumulative risk on CA/N report, independent of the effect of the number of visits and to determine the indirect effect of risk on CA/N report through the number of postpartum visits a family receives. However, it also likely that the number of visits is driven by the level of risk of the family. This model assesses the effect of cumulative risk on CA/N both directly and indirectly through the number of visits. This mediation model was tested separately for families initiating services in the postpartum and the prenatal period.

Table 16 provides the results of the mediation analysis for risk, postpartum visits, and later CA/N report. Analyses were completed separately for those who began services in the postpartum and prenatal period. The first two columns show the reduction in the estimate for the relationship between visits and later CA/N report before and after controlling for risk. The last three columns provide the total, direct, and indirect effect of risk when mediated through the number of visits. This analysis demonstrates that a large portion of the relationship between visits and later reports can be explained by the fact that high risk families receive more visits and also have a greater risk for report.

Table 16
Mediation Analysis for Risk, Number of Visits and CA/N Report

	Postpartum Visits	Visits Controlling for Risk	Mediation Effect of Risk Through Visits		
			Total	Direct	Indirect
Postpartum	0.22*	0.13*	0.18	0.18	0.01
Prenatal	0.69*	0.57	0.24	0.22	0.03

Note: *significant at $p < .05$

Aim 1 Results Summary. The firm aim of this study was to characterize the families receiving services by NFN within the current sampling frame. Further, this aim attempted to analyze service utilization and determine to what extent an increase in visits was associated with CA/N outcomes. Overall, this sample of families is a high-risk group of caregivers and children. Engaging with these families and retaining them in multiple visits over time was clearly a challenge. While there is not a specified number of visits that this home visiting model is hoping to achieve, there is some indication in the literature that multiple visits that span a wide range of the prenatal and postpartum period is ideal for supporting caregivers. Overall, 50% of families dropped out of services by their fourth home visit. For those receiving both prenatal and postnatal home visits, the total number of visits was higher.

Based on models predicting engagement and retention, individual factors at the caregiver and child level play a much larger role than nurse or geography in predicting these outcomes. Specifically, the number of visits a family receives appears to be strongly related to level of risk. Families that are higher risk and have more immediate concerns regarding the health of the infant and their needs as caregiver receive more home visits over a longer period of time. Families that report a higher level of caregiver stress using a validated screening tool stay engaged with services longer as well. These findings together indicate that perhaps the main factor driving level of service use is a caregiver's perception of short-term risk for their family.

Families who received more services also had a documented increase in the number of teaching skills that were provided by the nurse.

The level of service use did not appear to have a positive association with risk for later child maltreatment reports for this sample overall. In fact, families who engaged in services later and consumed more services in the postpartum period were at greater risk for later CA/N reports. However, those who received more visits prenatally were at a much lower risk for reports. Although statistical models attempt to control for level of risk and specific risk factors, there are likely other factors that are not captured in this analysis. The families who choose to receive home visits prenatally and those who begin services in the postpartum period are likely different types of families. This difference may be in terms of level of motivation of the caregiver, but there are no variables that measure this construct. There was an observable protective effect for a higher dose of services for caregiver who had a prior low birthweight baby and for those with a prior report of CA/N.

Results of Aim 2: Subgroup Analysis

Aim 2: To compare maltreatment reports and child health outcomes among those served in key policy-relevant subgroups.

This research aim was accomplished by examining the relationship between a selected set of policy-relevant subgroup to determine whether or not there were measureable differences in child maltreatment, child development, and caregiver mental health outcomes. Multivariate analyses were used to determine which demographic groupings were at greater risk for adverse outcomes. The following research questions address each of the three outcome areas separately examining results across the same subgroups.

Q2.1: Are there differences in rates of child maltreatment reports for family subgroups? The first subgroup analysis examined rates of child maltreatment for families across the subgroups. First a bivariate analysis was completed comparing raw rates of CA/N, then a multivariate Cox regression model was estimated to determine the adjusted hazard ratios for a CA/N event controlling for other subgroup characteristics, clustering at the nurse level, and time to event. Table 17 provides the results of the bivariate analyses predicting a later CA/N report and the results of the multivariate Cox regression model. The results are presented for any later CA/N report and for the more stringent, "qualified" CA/N report. These analyses were conducted using sample B, those with record of the child birth date.

Overall, 15.5% of the sample had a later report. The risk for report was higher at the bivariate and multivariate level for multiparous mothers, those initiating services postpartum, mothers with mental health, substance abuse or co-occurring disorders, rural families, and Biracial caregivers. Medically-fragile status and maternal age were significant at the bivariate level but were no longer significant in the multivariate model. Caucasians had a significantly higher rate of CA/N compared to African-Americans at the bivariate level, but when controlling for other factors, namely geographic location, the race effect was no longer significant for Caucasians. Caregivers and infant that were identified as Biracial had the highest rate of later report.

The largest effects in this analysis in terms of chi-square value at the bivariate and HR at the multivariate level were timing of the first visit and maternal behavioral health status. Only 1.1% of families who began services at the prenatal period had a later report, compared to 18.8% of those beginning in the postpartum period. In the multivariate model, this is associated with a HR = 18.4 (8.9 – 37.9) for risk of report for those who begin during the postpartum period

compared to prenatal. Compared to caregivers with no evidence of mental health or substance abuse concerns, those with co-occurring mental health and substance abuse have double the hazard for a later report.

Table 17
Subgroup Bivariate Analysis Predicting a CA/N Report Following NFN Services and Results of Multivariate Cox Regression Analysis Predicting Time to First Report

	% with CA/N Report			% with qualified CA/N report			Adjusted HR for CA/N Report		
	%	χ^2	*p*	%	χ^2	*p*	HR	HR 95% CI	
Full Sample Statistic	15.5			9.8					
Parity									
Multiparous (n=1509)	19.1	30.7	<.0001	12.5	25.1	<.0001	**1.52**	**1.24**	**1.86**
Primiparous (n=2111)	12.1			7.4					
First Visit Timing									
Prenatal (n=955)	1.1	137.6	<.0001	0.8	79.4	<.0001	**0.05**	**0.03**	**0.11**
Postpartum (n=2665)	18.8			11.9					
Newborn Health Status									
Medically-Fragile (n=1246)	19.2	25.9	<.0001	8.3	11.5	0.0	1.18	0.98	1.42
Healthy (n=2374)	12.7			11.8					
Maternal Behavioral Health		64.3	<.0001		42.6	<.0001			
Mental Health Only (n=530)	19.1			12.6			**1.63**	**1.29**	**2.07**
Substance Use Only (n=398)	23.5			15.9			**1.72**	**1.23**	**2.40**
Co-occurring (n=193)	27.0			15.7			**2.08**	**1.43**	**3.03**
No MH or SA (n=2499)	12.0			7.4			1.00		
Maternal Age		7.3	0.1		7.0	0.1			
<20 (n=1305)	13.1			8.1			1.00		
20-29 (n=1886)	15.9			10.1			0.87	0.73	1.04
30-34 (n=286)	18.0			12.7			0.86	0.61	1.20
35+ (n=143)	17.8			10.4			0.78	0.52	1.17
Family Geographic Location		21.0	<.0001		20.9	<.0001			
Urban (n=2816)	13.7			8.5			1.00		
Rural (n=499)	21.8			15.2			**1.56**	**1.12**	**2.17**
Suburban (n=291)	16.5			9.0			1.02	0.78	1.35
Caregiver Race		40.2	<.0001		49.1	<.0001			
Black (n=1700)	11.1			6.0			1.00		
Caucasian (n=1810)	18.6			12.8			1.21	0.98	1.50
Biracial (n=52)	25.0			16.7			**2.28**	**1.23**	**4.22**

Asian/AI/PI (n=30)	20.0			0.0			0.79	0.18	3.41
Child Race		46.3	<.0001		51.9	<.0001			
African-American (n=1564)	11.0			5.9					
Caucasian (n=1563)	18.5			12.6					
Biracial (n=225)	23.3			16.1					
Asian/AI/PI (n=28)	7.1			3.6					

Note: Adjusted HR controlling for maternal and child demographic variables, service variables, and clustering by nurse. HR in bold significant at $p < .05$

Q2.2: Are there differences in maternal mental health indicators for family subgroups? Maternal mental health indicators consisted of examining which caregivers were below above the cutoff for postpartum depression as indicated by a score of >13 on the EPDS for clinical range depression, >10 for problem range depression, and above the cutoff for high stress as indicated by a score above the sample mean for the ESI. Results of subgroups analysis for caregiver mental health are presented in Table 18. There was some consistency in the subgroups that predicted higher caregiver stress and higher postpartum depression. First, the caregiver who began their first home visits during the prenatal period were more likely to be above the cutoff for both stress and depression compared to those who began visits during the postpartum period. Those with mental health diagnoses and co-occurring disorder were higher for both stress and depression while those with isolated substance abuse disorders had a higher risk for depression but not stress.

There were other subgroups that had significantly higher risk for stress alone. The multiparous group, or mothers with prior pregnancies, was more likely to have higher stress. Rural caregivers and African-American caregivers were more likely to be above the threshold for high stress compared to their urban and Caucasian counterparts. These multivariate findings generally follow the bivariate findings. The only subgroup that did not have higher significantly

higher rates of stress or depression was caregivers of medically-fragile newborns compared to healthy newborns.

Table 18

Maternal Mental Health Outcomes across Identified Subgroups and Multivariate Models Predicting Problem Stress and Postpartum Depression

	Adjusted OR for EPDS in clinical range (n=1561)			Adjusted OR for "High Stress" ESI (n=1889)			% "High Stress" ESI (n=1962)		EPDS Clinical Range (n=1620)		EPDS Problem Range (n=1620)	
	OR	OR 95% CI		OR	OR 95% CI		%	*p*	%	*p*	%	*p*
Full Sample Statistic							38.1		7.5		11.9	
Parity												
Primiparous (n=2111)	1.00			1.00			32.0		5.8		9.5	
Multiparous (n=1509)	1.54	0.98	2.41	**1.76**	**1.39**	**2.24**	46.5	*	9.9	*	15.4	*
First Visit Timing												
Postnatal (n=2665)	1.00			1.00			33.3		6.1		10.0	
Prenatal (n=955)	**1.76**	**1.16**	**2.69**	**1.81**	**1.43**	**2.29**	50.2	*	10.9	*	16.7	*
Newborn Health Status												
Healthy (n=2374)	1.00			1.00			38.2		7.2		11.3	
Medically-Fragile (n=1246)	1.08	0.70	1.66	0.94	0.74	1.18	37.8	ns	8.0	ns	13.1	ns
Maternal Behavioral Health								*		*		*
No MH or SA (n=2499)	1.00			1.00			32.4		3.8		6.6	
Mental Health Only (n=530)	**5.78**	**3.64**	**9.18**	**2.75**	**2.05**	**3.68**	55.0		18.4		27.8	
Substance Use Only (n=398)	**2.22**	**1.18**	**4.20**	1.27	0.89	1.80	36.6		8.3		13.0	
Co-occurring (n=193)	**5.10**	**2.58**	**10.07**	**3.41**	**2.13**	**5.45**	59.3		16.3		25.0	
Maternal Age								*		ns		ns
<20 (n=1305)	1.00			1.00			33.9		5.7		9.9	
20-29 (n=1886)	1.07	0.63	1.82	1.07	0.83	1.37	40.7		8.6		12.8	
30-34 (n=286)	1.20	0.53	2.75	0.83	0.53	1.29	38.5		9.6		16.2	
35+ (n=143)	0.46	0.12	1.73	0.86	0.48	1.55	37.8		4.5		9.0	
Family Geographic Location								*		ns		*
Urban (n=2816)	1.00			1.00			37.1		7.1		11.0	

Rural (n=499)	1.63	0.90	2.93	**2.49**	**1.76**	**3.54**	46.9		10.3		18.7	
Suburban (n=291)	0.90	0.39	2.06	1.12	0.69	1.81	32.3		6.7		10.1	
Caregiver Race								*		ns		ns
Caucasian (n=1810)	1.00			1.00			31.9		7.2		12.0	
Black (n=1700)	1.48	0.92	2.37	**2.40**	**1.85**	**3.11**	44.2		7.9		11.6	
Biracial (n=52)	.			.			50.0		8.7		21.7	
Asian/AI/PI (n=30)	.			.			5.0		0.0		6.7	

Note: Adjusted estimates are controlling for maternal and child demographic variables, service variables, and clustering by nurse. Estimates in bold are significant at $p < .05$, For Bivariate tests, $*p < .05$

Q2.3: Are there differences in child developmental health indicators for family subgroups? Child development was assessed using the ASQ and the cutoff scores across the six domains of developmental concerns. Families received the ASQ screen if they were still receiving home visits by the time the child was four months old. There were 708 infants screened at 4 months and 85 infants that were screened at 18 months. Overall, 989 families received an ASQ screen at any given age. The results reported were for families that were below the cutoff point at any age across the 4, 6, 8, 10, 12, 14, 16, or 18 month screener. Children were assessed across six developmental concerns and if they were below the cutoff for any of these domains, they were coded positive for "any" concern. An OR was developed based on a multivariate model predicting the risk for each group to have any ASQ concern.

The medically-fragile newborn group was clearly the highest risk group for developmental concerns. This group had three to six times the risk for any concern controlling for other variables in the model. At the bivariate level, the medically-fragile group had a higher rate of concerns across all six domains compared to healthy children. After controlling for this grouping, none of the other subgroups had an adjusted odds ratio over one indicating higher risk for developmental problems.

At the bivariate level, there were some subgroups that had higher rates of developmental concerns. Multiparous mothers had children with a higher rate of any developmental concerns and higher rates for each domain except problem-solving. Mothers who initiated services postnatally and those with mental health or substance abuse problem had children with higher rates of gross motor concerns. Mothers over the age of 35 had children with more communication concerns. There were no significant differences in developmental concerns across geographic subgroups or race of the child.

Table 19
Results of Subgroup Analysis for ASQ Developmental Screening Tool

	Adjusted OR for ASQ Any Concern			ASQ Any Concern		Commun-ication		Gross Motor		Fine Motor		Problem Solving		Personal-Social	
	OR	OR 95% CI		%	*p*	%	*p*	%	*p*	%	*p*	%	*p*	%	*p*
Full Sample Statistic				21.9		7.2		7.4		11.4		11.6		9.4	
Parity															
Primiparous (n=2111)	1.00			17.8		5.1		5.7		8.3		10.3		7.1	
Multiparous (n=1509)	1.42	0.96	2.10	26.6	*	9.6	*	9.4	*	15.0	*	13.1	ns	12.1	*
First Visit Timing															
Postnatal (n=2665)	1.00			22.9		7.9		8.6		12.4		12.6		10.2	
Prenatal (n=955)	0.98	0.65	1.48	18.9	ns	5.0	ns	3.8	*	8.4	ns	8.8	ns	7.1	ns
Newborn Health Status															
Healthy (n=2374)	1.00			12.3		3.2		3.9		6.3		5.4		4.5	
Medically-Fragile (n=1246)	**4.26**	**2.97**	**6.11**	36.5	*	13.2	*	12.6	*	19.2	*	21.2	*	17.0	*
Maternal Behavioral Health					ns		ns		*		ns		ns		ns
Mental Health Only (n=530)	1.29	0.82	2.01	24.2		7.0		8.4		11.1		11.6		11.1	
Substance Use Only (n=398)	0.99	0.57	1.71	28.2		8.4		13.5		12.6		11.5		13.5	
Co-occurring (n=193)	0.78	0.37	1.66	23.1		6.3		0.0		10.4		5.8		5.8	
No MH or SA (n=2499)	1.00			20.0		5.8		6.7		11.1		12.1		8.6	
Maternal Age					*		*		ns		ns		ns		ns
<20 (n=1305)	1.00			16.9		3.6		5.8		8.8		11.4		7.1	
20-29 (n=1886)	1.14	0.74	1.77	24.0		8.8		7.4		13.0		11.8		10.4	
30-34 (n=286)	1.05	0.51	2.13	23.8		6.3		12.5		7.5		8.8		10.0	
35+ (n=143)	1.14	0.46	2.86	31.4		17.1		8.6		20.0		17.1		14.3	
Family Geographic Location					ns		ns		ns		ns		ns		ns

Urban (n=2816)	1.00			21.5		6.4		6.5		11.5	10.6	9.2
Rural (n=499)	1.10	0.69	1.75	23.4		9.1		10.1		10.1	14.4	10.5
Suburban (n=291)	0.86	0.41	1.77	19.4		7.5		7.5		13.4	11.9	9.0
Child Race					ns		ns		ns		ns	ns
African-American (n=1564)	1.00			21.0		6.2		6.2		12.0	11.1	9.1
Caucasian (n=1563)	1.09	0.71	1.66	22.3		8.1		8.5		9.8	12.5	9.5
Biracial (n=225)	1.39	0.71	2.70	25.4		7.0		5.6		18.3	8.5	14.3

Note: Adjusted estimates are controlling for maternal and child demographic variables, service variables, and clustering by nurse. Estimates in bold are significant at $p < .05$, For Bivariate tests, *$p < .05$

Aim 2 results summary. The results of Aim 2 indicate that there are significant differences in outcomes across important subgroups served by NFN. The agency provides home visits to a diverse pool of families with a wide set of risk factors and life situations. This study identified seven different subgroups of families to explore child maltreatment, child development, and maternal mental health. These groups were multiparous versus primiparous mothers, those who began services prenatally versus postnatally, medically-fragile newborns versus healthy newborns, caregivers with mental health and substance abuse issues versus those without behavioral health issues, different caregiver age groups, family geographic locations, and child racial groups. When examining the results across outcomes that span both child and maternal health, there are several themes that emerge.

First, first-time mothers and mothers who already have children appear to have quite different levels of risk for all outcomes explored. Specifically, first-time mothers are at lower risk for child maltreatment, child developmental concerns, and maternal stress even when controlling for a host of other factors. This finding directly pertains to the policies that expand home visiting programs that only target first-time mothers and exclude mothers with children from services. Second, there appears to be a very real difference in this population of families who begin services during the prenatal period and those who begin services after the baby is born. This difference can be seen in terms of risk for child maltreatment, depression, stress, and gross motor child development. Prenatally referred women reported higher depressive symptoms and caregiver stress yet had much lower rates of child maltreatment and had lower rates of ASQ concerns across all domains. This may be a difference in the types of caregivers who seek out prenatal services compared to those who seek out postnatal services, but it also may indicate some protective effect of prenatal visits.

Finally, this subgroup analysis demonstrated the impact of mental health and substance abuse issues on the mother as well as on the child. Women with behavioral health concerns had higher reported rates of depressive symptoms and caregiver stress and also had around twice the rate of child abuse and neglect reports. This is a particularly vulnerable group of families both for the impact on the caregiver but also for the additional risk placed on the infant.

Results of Aim 3: Program Effectiveness

Aim 3: To compare subsequent child maltreatment reports for families that receive NFN services to a matched comparison group using a quasi-experimental design utilizing propensity score and survival analysis.

The purpose of this final research aim is to attempt to isolate an estimate for the treatment effect of NFN services for preventing child maltreatment. An effect for the entire sample is examined first and then separately for the postpartum referred group. In the absence of randomization, quasi-experimental techniques using a dropout comparison group were used. Using the counterfactual framework, the analysis presented is based on the premise that a reasonable comparison can be made if selection bias can be adequately reduced between a family that engages in more than one visit and a family that drops out after one visit. The propensity score was defined as the conditional probability that each family receives more than one home visit based on a set of pre-treatment variables. After the propensity score (PS) was estimated, several conditioning methods were used to eliminate this bias associated with dropout status including regression adjustment, stratification, and propensity score matching. This section will first provide more detail regarding CA/N reporting outcomes of the service sample at the bivariate level. Then, the results of PS estimation process and conditioning models will be reported for models estimating the likelihood of a report following services. Finally, results from

survival analyses including Cox regression models will be reported that consider the timing of services and reports in predicting the time to a first report of maltreatment.

Q3.1: To what extent can any differences in the rate of maltreatment reports be attributed to the Nurses for Newborns intervention? The following table displays the percent of families that had a CA/N report across dimensions of service population, treatment group, and the timing of the report. The service population is the original risk group categorization for each family, the treatment group is whether the family is in the services treatment condition or in the dropout comparison group condition for the final PS matching analysis, and the timing of the services categorizes whether the report occurred before, during, or after NFN services. Table 20 presents the results of chi-square tests of association comparing treatment and dropout condition for each service population and time period. Rates with an asterisk represent a significant association at $p < .05$ for that contrast. This analysis represents those families in sample B, or those that had a birth record for match (n = 3376). In the total sample, the dropout comparison group was more likely to have a CA/N report prior to services but less likely to have a report between referral and termination of services. There were no significant differences for reports after services.

Table 20
Rate of CA/N Report by Timing of Report, Service Population, and Treatment Group

Service Population	Treatment Group	n	Any Report	Prior to Services	During Services	After Services
Total Sample	T	2750	19.4	1.9*	4.6*	15.4
	C	626	16.5	3.7	2.1	13.9
Poverty	T	983	12.6	1.1*	2.6	10.5
	C	258	11.2	3.5	1.2	8.9
At-Risk Baby	T	674	24.8	2.8	4.3	20.2
	C	178	24.7	5.6	2.8	20.2
At-Risk Mom	T	650	29.7*	2.5	9.9	22.2
	C	90	18.9	4.4	4.4	16.7
Teen Parent	T	443	10.8	1.1	1.8	9.0
	C	100	13.0	0.0	1.0	13.0

Note: *$p < .05$

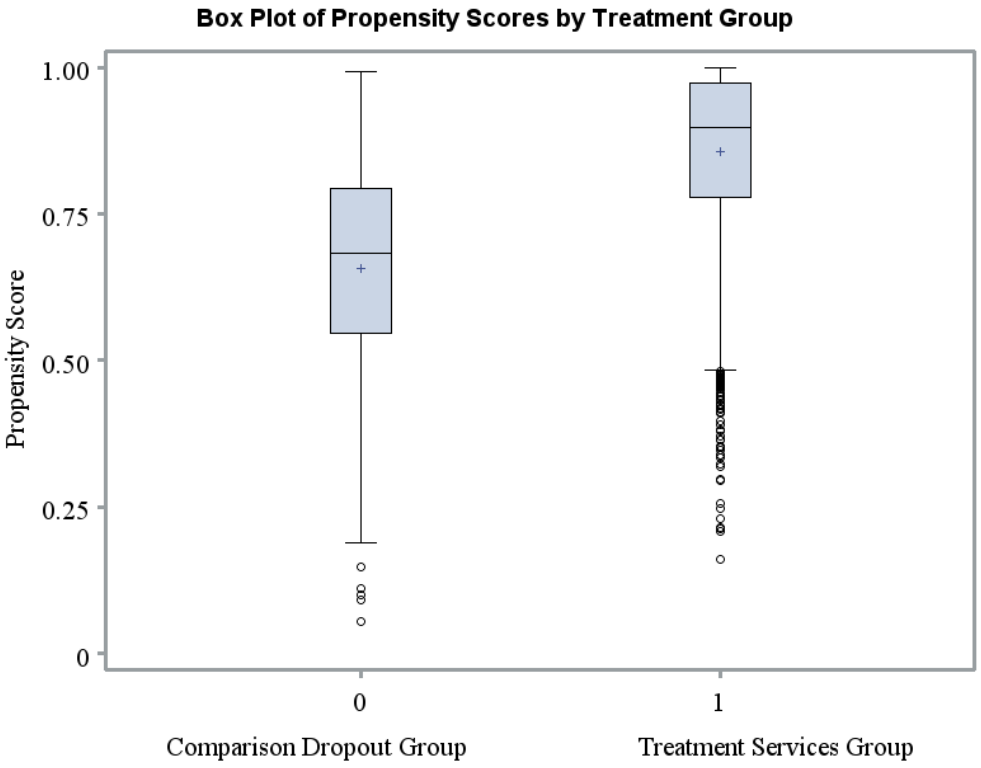
The analysis in Aim 1 explored differences between families in their engagement status. These models found many key differences between families that initially engage in services and those that drop out. For example, the engagement findings reported that families with a higher level of cumulative risk were more likely to initially engage with services. A simple comparison of the raw rates of CA/N between the treatment and dropout comparison group does not incorporate this bias between the two groups. We also know that cumulative risk and multiple stressors likely increase the risk for the outcome, later CA/N report. So, the comparison in the raw rates is biased. By estimating the PS and conditioning the analysis on these selection variables, a more adequate comparison can be made.

Propensity score estimation. Sample C, those families with a birth date and at least one postpartum visit were used for the propensity score analysis (n = 3299). The propensity score was estimated by regressing the dichotomous treatment group condition on a set of pre-treatment covariates. The model fit of the logistic regression model was assessed and the Hosmer-Lemeshow goodness-of-fit $\chi^2 = 7.23$, $p = .51$, the area under the receiver operator curve (ROC) plot was .80, and the pseudo $R^2 = .30$ indicating a good fit of the data predicting first visit

dropout propensity score. Figure 12 displays a boxplot of the propensity score by treatment group. One important consideration before conducting PS analysis, particularly matching, is to ensure there is adequate overlap between the treatment conditions on the PS. This boxplot indicates that there appears to be sufficient overlap between the two treatment conditions. The next figure displays the same information but in the form of a frequency distribution of the PS. This figure also displays the sample statistics for the PS across the two treatment conditions.

The fit for the model predicting the PS was adequate and there appeared to be sufficient overlap in scores for conditioning methods. It is also important to assess the reduction in bias that can be achieved using the PS. This can be done by comparing the means and standardized differences for key predictor variables before and after matching on the PS. The results provided here are for a variable optimal match with a caliper set at $d = .01$ for the matching width. More detail will be provided about the different matching techniques used, but this initial analysis provides an indication that the modeling is achieving the goal of bias reduction.

For this specific match, there was a complete reduction of bias in the propensity score and a 71% reduction in bias of the cumulative risk score. Other selected covariates have a decrease in the standardized differences and a high level of bias reduction. These are only a selection of covariates that were known to have bias prior to matching. The estimating equation contains many more variables and interaction effects than are listed. This table also provides the p-values for the comparison of the means for the treatment and comparison condition before and after matching. Before matching, all of the selected covariates were significantly different in the two groups and after, there were no significant differences. Given that the outcome measure is likely related to many of these baseline characteristics, a direct comparison between these two groups would be biased and invalid.



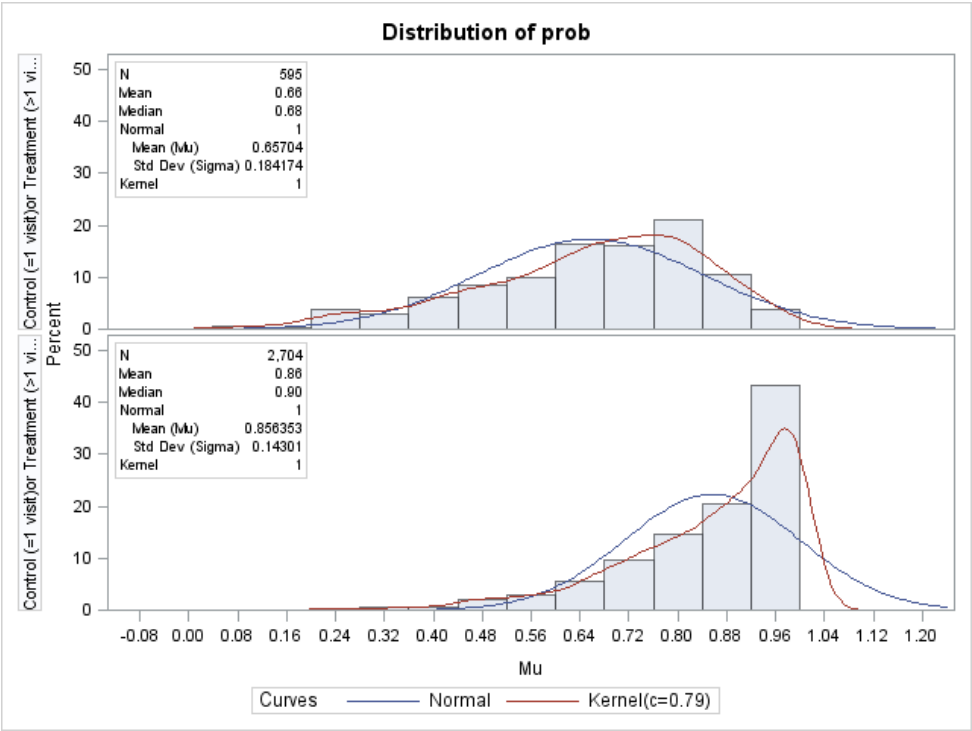


Figure 12: Distribution of Propensity Score for Dropout and Treatment Groups

Table 21
Comparisons of Means, Standardized Differences, and Bias Reduction For Matched Sample

	Full Sample (n=3299)				Matched Sample (=1102)				
	C (n=595)	T (n=2704)	*p*	*d*	C (n=551)	T (n=551)	*p*	*d*	% Reduction in Bias
Logit of Propensity Score	0.66	0.70	*	1.183	0.66	0.66	ns	0.001	100.0%
Cumulative Risk Score	3.56	4.07	*	0.249	3.54	3.39	ns	0.075	71.1%
High-Risk Mom Subgroup	0.14	0.24	*	0.244	0.14	0.13	ns	0.065	88.6%
Prenatal Referral	0.08	0.29	*	0.121	0.09	0.09	ns	0.055	100.0%
Caregiver Current DV	0.03	0.06	*	0.121	0.03	0.02	ns	0.023	76.0%
Caregiver Prior CA/N	0.01	0.05	*	0.209	0.01	0.01	ns	0.014	85.4%
Rural Zip	0.12	0.15	*	0.083	0.12	0.11	ns	0.017	72.9%
Baby Age at Referral	2.18	1.75	*	0.258	2.14	2.11	ns	0.035	94.4%
Caregiver Education Level	1.02	0.93	*	0.075	1.03	1.02	ns	0.004	90.1%
Zip % Child Poverty	20.25	22.05	*	0.073	20.33	20.21	ns	0.014	93.5%
Median Family Income	4.43	4.27	*	0.073	4.41	4.43	ns	0.005	90.0%

Note: d = standardized differences, Percentage bias reduction is calculated by (1- Di)/Dj *100% where Di and Dj are group difference in covariates means,
*$p < .05$

PS regression adjustment. The first conditioning method is a standard logistic regression model using the propensity score and the treatment condition to predict the outcome of CA/N report. For this dichotomous outcome, the treatment effect is considered the adjusted odds ratio of the group variable. The estimate for this analysis was $b = .32$, $p = .02$, which produced an OR = 1.38 95% CI (1.05 – 1.83). After controlling for the propensity score, the

treatment group had a 5-83% greater risk of having a CA/N report compared to the dropout comparison group.

PS stratification. The sample was stratified into quintiles based on the propensity score and the treatment effect was determined for each subclass. There were 660 families in each PS quintile. Table 22 provides the range of propensity scores and results for the PS stratification analysis. When examined across quintiles, there was not a significant treatment effect for any group. For this reason, there was no need to combine the treatment effect across the strata.

Table 22
Result of Propensity Score Stratification Predicting a later CA/N Report for Logistic Regression and Survival Analysis Cox Regression

PS Quintile	PS Range	OR	95% Lower	95% Upper	HR	95% Lower	95% Upper
1	.05 - .69	1.33	0.89	1.96	1.30	0.91	1.85
2	.69 - .82	1.54	0.90	2.63	1.54	0.93	2.54
3	.82 - .90	1.43	0.73	2.79	1.47	0.79	2.73
4	.90 - .97	1.32	0.45	3.88	1.23	0.45	3.32
5	.97 - .99	0.25	0.03	2.42	0.24	0.03	1.73

PS Matching. Propensity score matching is not a single uniform analytic technique but has different options with regarding to the matching algorithm, selection process, and caliper size. While the varying methods were used, the general principle is that the observation pairs are selected that minimize the distance between the propensity scores. The sensitivity analysis section at the end compares results across these different combinations of matching methods. All of the matching results reported in this section and the time-to-event analysis uses 1:1 variable optimal matching without replacement. The results of the propensity score matching techniques yielded results similar to the regression adjustment and stratification findings. The estimation of the overall treatment effect was not significant, OR = 1.34 95% CI (.97 – 1.84) with a trend in the direction of greater risk for report for families that received more than one

visit. A caveat is important here, related to the fact that although groups were balanced according to prenatal engagement, the prior analyses indicate that this is actually a very different subpopulation. Unfortunately, the rate of maltreatment reporting was so rare, that it was not possible to use PS techniques to develop two separate groups.

Q3.2: To what extent can any differences in the timing of maltreatment reports be attributed to the Nurses for Newborns intervention? Results for this section repeat the research question but incorporate time-to-event survival analytic methods. Since families enter and terminate from the program at different times, survival analyses were conducted to model the time at risk for the outcome of CA/N report.

Bivariate survival analysis. Since families enter and terminate from the program at different times, survival analyses were conducted to model the time at risk for the outcome of CA/N report. The following figure presents a survival curve for the product-limit estimate of the survivor function across service populations. This purpose of this figure is to provide an orienting baseline level of risk for CA/N report across the different types of families involved in these services. The values at the bottom inset of the figure provide the number of families that are "at-risk" at the beginning of the time (their program termination date) and after one (365 days), two (730 days), and three years (1095 days) following termination of services. The log-rank test of equality is significant ($\chi^2 = 85.5$, $p < .0001$). The Sidak multiple-comparison adjustment was used to conduct paired test of each level of the service population strata. This analysis essentially confirmed what is clear from visual inspection of the survival curves, the poverty and teen parents groups (top curves) are similar (Sidak $p = .10$) and the at-risk baby and at-risk mother groups (bottom curves) are statistically similar (Sidak $p = .99$). Comparisons between any group between the top and bottom sets of curves are significant at Sidak $p < .0001$.

This finding confirms that families in either of the two identified higher risk groups have greater risk for a CA/N later report compared to the lower risk service populations. The next set of analyses will attempt to use propensity score analyses to isolate a treatment effect of service use for all families in the study.

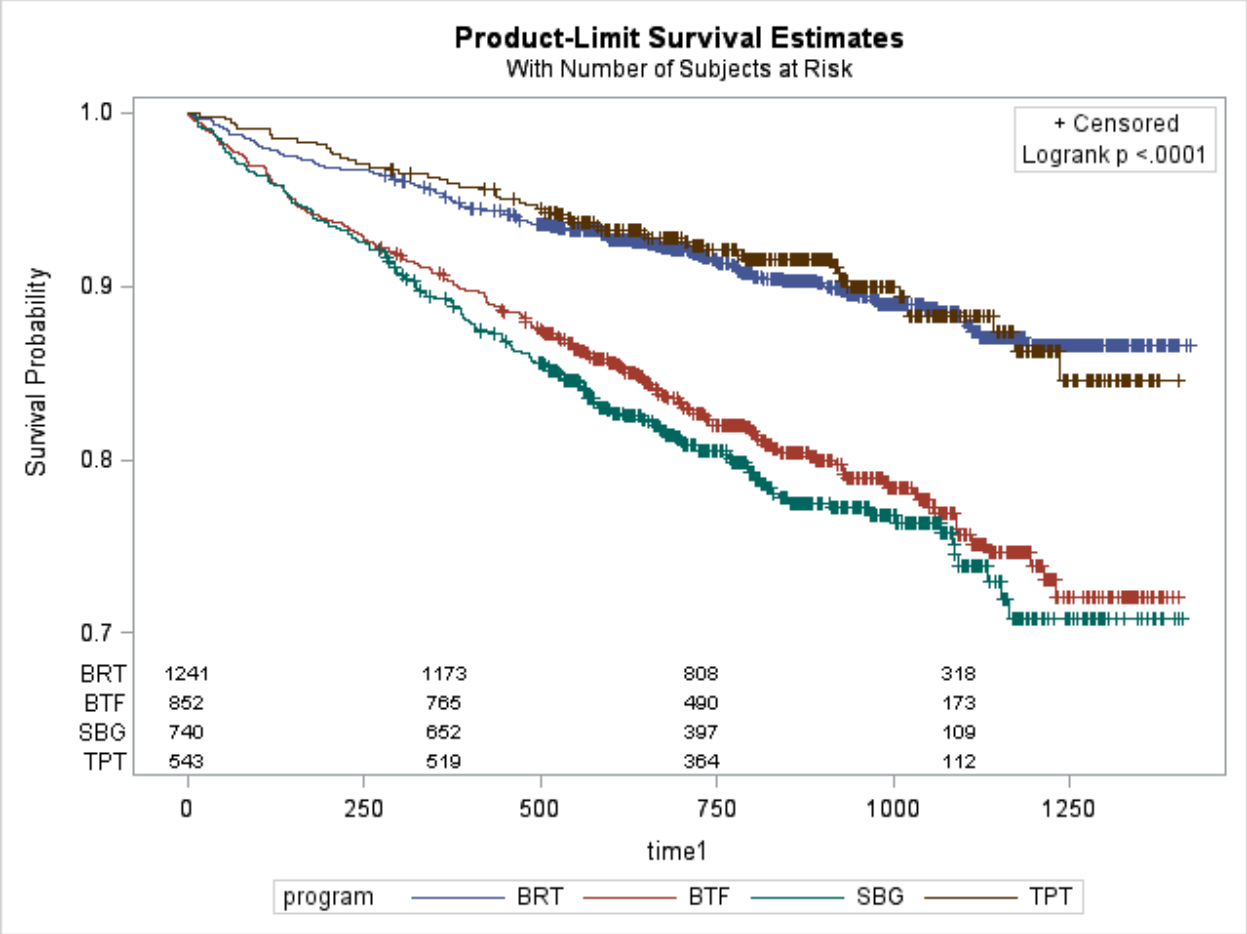


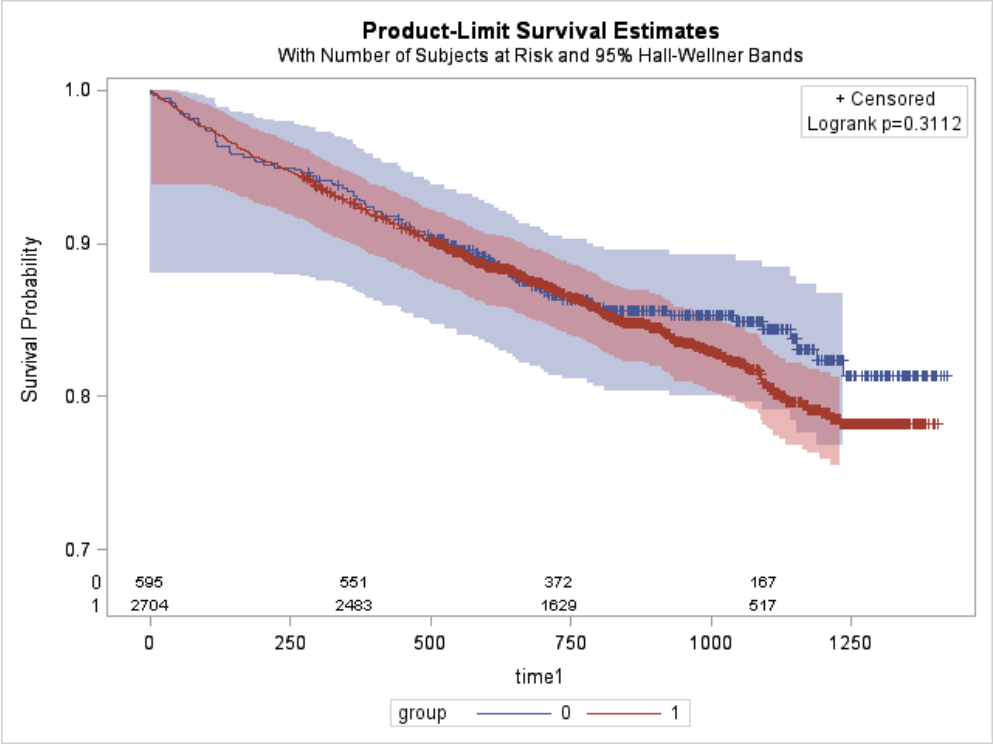
Figure 13: Survival curve for time to event of later CA/N report across service populations

Survival PS regression adjustment. When considering CA/N report as a time to event outcome with the propensity score and treatment group variable as the covariates, the estimate for treatment group was $b = .33$, $p = .01$, which was associated with a HR = 1.39, 95% CI (1.07 – 1.79). This outcome can be interpreted as an increase in 39% of the hazard rate for a report

following NFN services for families who receive more than one visit compared to those who drop out after the first visit after controlling for the propensity score.

Survival PS stratification. The results of PS stratification for the survival analysis are included in Table 20. The results were essentially identical to the bivariate results for any report regardless of time. The 95% CI for the HR of all groups spanned 1 and were not-significant.

Survival PS matching. The results of PS matching were similar to the bivariate results when considering the time-to-event in the analysis. The following figures provide survival curves for the full sample first and then the matched sample. In both cases the 95% Hall-Wellner Bands for the survival curves overlap and the log-rank $p > .05$ indicating equivalency of the hazard rate over time for the treatment and comparison conditions. The blue line represents the comparison group and the red line represents the treatment group.



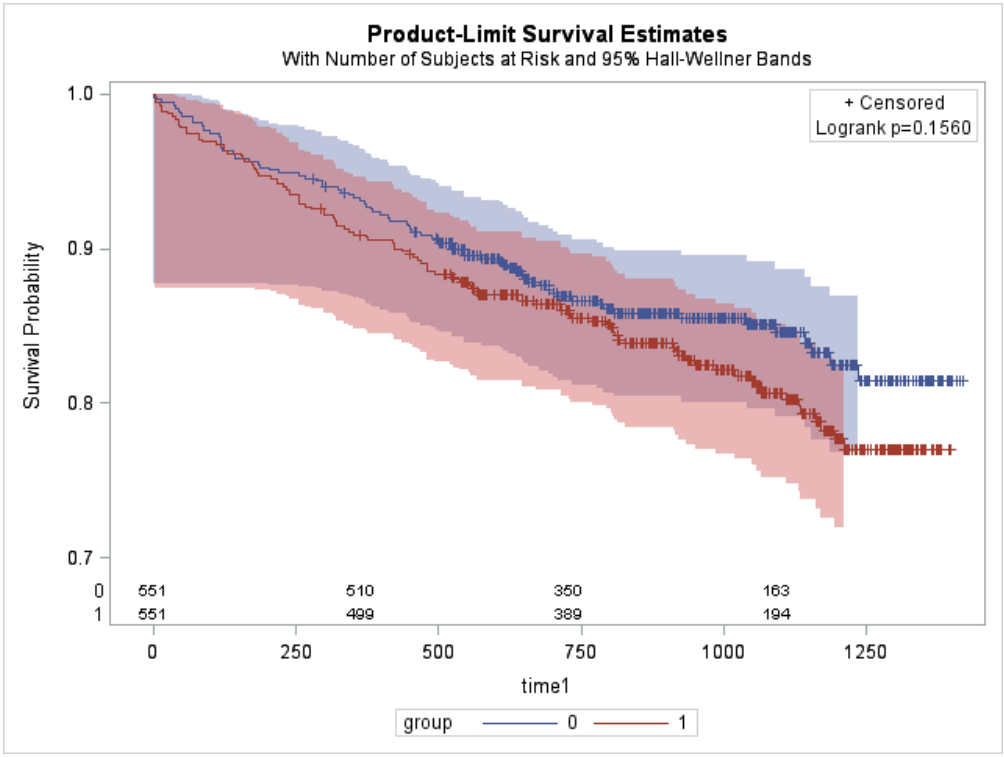


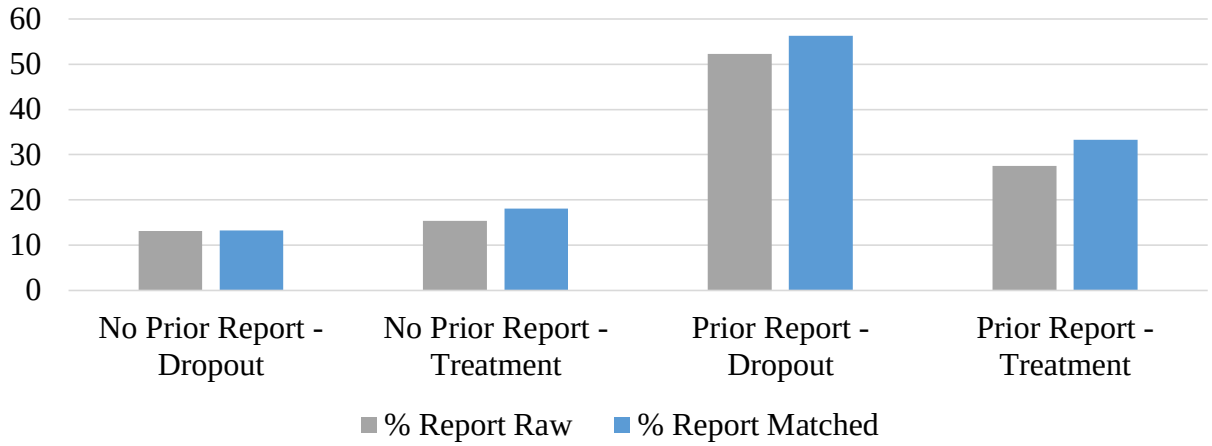
Figure 14: Survival Curves Before and After PS Matching Comparing Treatment and Dropout Comparison Group Risk for CA/N Report

Subgroup analysis. The PS matching methods indicated that there was not a significant treatment effect for the entire sample. Subgroup analyses were conducted that included moderation interaction effects to determine if the treatment effect differed across specific subpopulations. First, the policy-relevant subgroups analyzed in Aim 2 were included in a model as well as those groups determined to be at differing risk for a later CA/N report in the Aim 1 moderation analysis. There were no significant interaction effects for service population, race, cumulative risk, parity, infant health status, caregiver behavioral health, or caregiver age. There was a statistically significant for prior CA/N report and when race was removed from the model there was a significant effect for geographic location. These results will be explored in more detail below.

Children with a prior CA/N report have over twice the hazard rate for a later report (HR=2.44, Wald 95% CI 1.28 – 4.63) when controlling for other factors in the model and nurse clustering. However, there was also a significant interaction term by comparison group ($B_{priorrep \times group} = -1.71$, $p < .001$) in the matched sample. To ease interpretation of the interaction term, the bivariate results for the full and matched sample comparing showing the relationship between group and later report based on prior report status is shown in the following Figure 15. The key to interpreting this figure is to examine the difference between dropout and treatment group for the two no prior report groups on left and the difference for the two prior report groups on the right. The light and dark bars are simply the % for the raw and matched samples to demonstrate that the finding is not a relic of the matching process. While there is no difference between treatment and control for those with no prior reports, there appears to be a large protective effect for the treatment condition among families with a prior report. Over half of families with a prior

report that dropout have a later CA/N report while those who engage in services for at least one more visit have about half the report rate.

Figure 15. Rates of Later CA/N Report by Treatment Group Comparing those with Prior CA/N Report

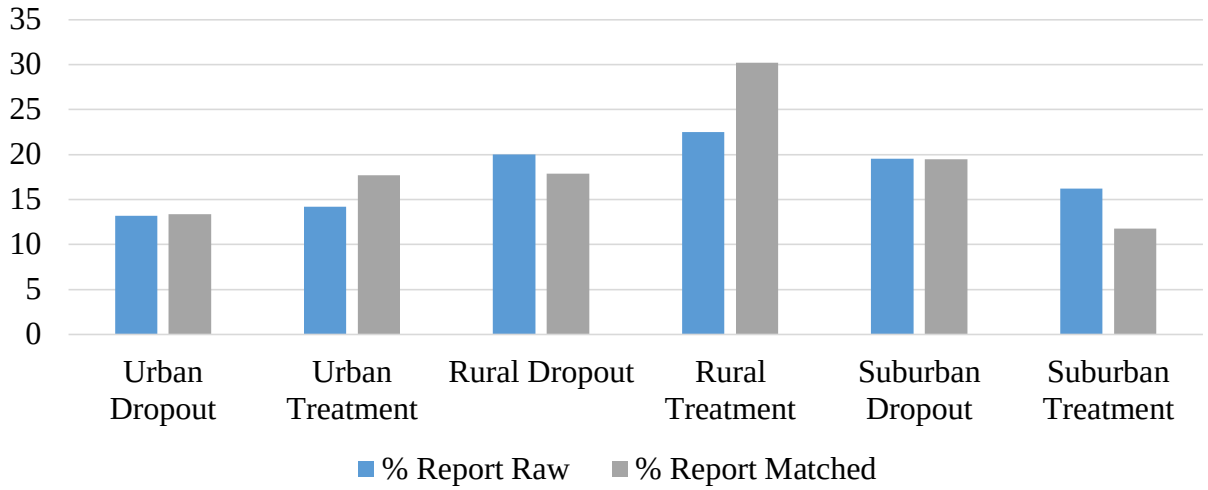


While the interaction effect indicates that the relationship between treatment and outcome is different for this group, a post hoc analysis is needed to examine an average treatment effect for this subgroup. The number of families with later reports was relatively low in the sample, so to further explore the treatment effect a post hoc analysis was conducted using propensity score regression adjustment for the 74 families with a prior CA/N report. This analysis found a main treatment effect of B_{group} -1.11, $p = .06$ which corresponds to a HR = 0.33, 95% CI 0.11 – 1.04 suggesting 94% confidence that there is a true average treatment effect for families with a prior report.

The second significant interaction effect was based on the family's geographic location based on their zip code categorized as rural, urban, or suburban. While there was not a main effect for geographic location predicting a later report, the interaction term was significant for urban geography ($B_{urban\ x\ group} = .86$, $p < .05$) and for rural geography ($B_{rural\ x\ group} = 1.09$, $p < .01$) compared to the suburban reference group. Additional analyses found that the difference

between rural and urban was the same. This finding suggests that the treatment effect for suburban families is different from urban and rural families. The figure below shows the frequencies of later report for the different geographic groups by treatment condition. While the rate of later report is higher for the treatment group for urban and rural families, the rate of CA/N is lower among suburban treatment families compared to dropout families.

Figure 16. Interaction Effect of Geography and Treatment Condition Predicting Follow-up CA/N Report



A post hoc subgroup analysis was also conducted for geographic groups to examine the treatment effect for each group using a regression adjustment propensity score method. This analysis determined that there when the suburban geographic group (n = 262) did not find a significant treatment effect B_{group} = .19, p = .65 indicating that the effect of group was lower for suburban families compared to other groups but was not a significant overall protective effect.

Postpartum subsample. Given the prior findings indicating differences in outcomes for families who initiate services prenatally and those who initiate during the postpartum period, an additional match was conducted limiting the analysis only to those who initiate in the postpartum period. The previous results were for families who had any visits in the postpartum period, but included those who initiated in the prenatal period. Although the matching process in this

previous analysis use first visit timing in the propensity score and analyses indicated that bias had been diminished for this factor, prior analysis suggests that these two groups should be examined separately. No propensity score analysis is conducted for the prenatal visit group separately as the rate of later report is too low for additional multivariate analysis, particularly ones that result in data reduction such as matching.

A propensity score model for postpartum-only families (n = 2665) was estimated and assessment of the balance after matching was assessed as described above. Results indicated a similar distribution of propensity scores and balance of pre-treatment covariates after matching as in the full-sample analysis. In the matched sample (n = 856), the OR for a later CA/N report was 1.52 and prior to matching the OR = 1.47, indicating a similar finding before and after matching. Prior to matching, 20.1% of the engaged group with more than one visit had a later report compared to 14.6% of the dropout group. After matching, 20.8% of the engaged treatment group had a later report compared to 14.7% of the dropout group.

A similar test was conducted to determine whether increasing the required dosage to be considered part of the "treatment" group had an impact on findings by limiting this group to those with four or more visits (n = 1255) and then to those with seven or more visits (n = 886). Findings indicate a similar direction but non-significant treatment effect for 4+ visits (OR = 1.34, 95% CI .89 to 2.10) and for 7+ visits (OR = 1.22, 95% CI .70 to 2.16). The rate of later report among the matched samples for higher doses of service was not significant and was 18.5% (vs. 14.3% comparison) for those with 4+ visits and was 21.7% (vs. 18.4% comparison) for those with 7+ visits. Overall, these findings are consistent with those reported in more detail for the sample including those who initiated services prenatally.

Sensitivity analysis. Given the flexibility and available options for conducting propensity score analyses, particularly propensity score matching, it is important to conduct a sensitivity analysis to ensure that similar effects are found regardless of the analytic technique selected. Several different propensity score conditioning methods and matching techniques were discussed in the methods section. This sensitivity analysis will examine the treatment effect (OR for treatment group predicting a later CA/N report) when using either greedy or optimal matching, using different matching replacement schemes, and adjusting the caliper width.

One consideration when selecting a matching method is the width of the caliper or the maximum allowable difference between propensity scores of matched observation. Given that there are fewer comparison (n = 595) families than there are treatment (n = 2704), the comparison families are the limiting selection group in 1:1 match without replacement. A smaller caliper increases the precision of the matching but also reduces the sample size. A width of .01 was selected initially and reduced the comparison group sample size to n = 551. There is no accepted method for assigning the caliper width but one suggestion from the literature is to use a caliper width .2 times the standard deviation of the logit of the propensity score (Rosenbaum & Rubin, 1985; Austin, 2011), or for this sample = .2(.037) = .0074. When this caliper width was used, it reduced the comparison sample size to n = 546. With a caliper width = .0001, which is a relatively narrow caliper, there were still 213 families in the comparison condition that found treatment matches. The sensitivity analysis compared caliper sizes = .01, .001, and .0001 widths. The sensitivity analysis confirms the results reported were consistent across different matching schemes, there was not a significant relationship between treatment group and risk for later CA/N report.

Table 23
Propensity Score Matching Sensitivity Analysis across Matching Methods Estimating a Later CA/N Report

Matching Type	Replacement	Caliper Size	Treatment Group Size	Comparison Group Size	Treatment Effect Estimate (95% CI)
Greedy NN	Y (2:1)	None	2704	595	1.02 (0.78 – 1.34)
Greedy NN	N	None	595	595	0.96 (0.68 – 1.34)
Greedy NN	Y (1:1)	0.1	2704	595	1.04 (0.79 – 1.36)
Greedy NN	N	0.001	455	455	0.77 (0.51 – 1.16)
Greedy NN	Y (2:1)	0.001	926	492	0.97 (0.70 – 1.36)
Variable Optimal	N	0.01	486	486	1.33 (0.92 – 1.92)
Variable Optimal	N	0.001	449	449	1.36 (0.93 – 2.01)
Variable Optimal	N	0.0001	213	213	1.45 (0.82 – 2.56)

Another question stemming from this analysis was the extent to which the treatment and comparison group represented truly different service conditions. This analysis hinges on the assumption that families who receive one visit only have a distinct service experience compared to those who receive more than one visit. Further, this analysis assumes that all of the families who receive at least one visit are receiving a relatively consistent treatment condition. However, prior analyses indicated that most families receive a fairly low level of home visits. In fact, over half of the families have dropped out by the fourth visit. This analysis is comparing families with one visit, largely to those with two, three, and four visits. This can be problematic in that there is not a sufficient distinction in the service experience to truly test a differential effect for a "treated" group.

The following table provides an analysis of the change in treatment effect if the minimum number of visits is increased for inclusion in the "treatment" group. The minimum for all prior analyses was two visits and this value was increased to three, five, seven, and 10 visits. Variable optimal matching was used for all of these analyses but the caliper size was adjusted to provide further sensitivity to the findings. While a significant treatment effect was not detected as the minimum visits were increased, the valence of the point estimate switched from an increased risk

for CA/N to a decreased risk. As the number of visits increases, the sample size available for matching also decreases which decreases the power to detect a significant difference.

Table 24
Propensity Score Matching Sensitivity Analysis for Different Minimum Visits for Treatment Group

Matching Type	Minimum Visits for Treatment Group	Caliper Size	Matched Group Size	Treatment Effect Point Estimate (95% CI)
Variable Optimal Matching	3	0.01	470	1.33 (0.92 – 1.95)
Variable Optimal Matching	3	0.001	183	1.31 (0.73 – 2.36)
Variable Optimal Matching	5	0.01	334	0.85 (0.54 – 1.34)
Variable Optimal Matching	5	0.001	104	1.43 (0.68 – 3.05)
Variable Optimal Matching	7	0.01	176	0.91 (0.48 – 1.68)
Variable Optimal Matching	7	0.001	76	0.63 (0.24 – 1.64)
Variable Optimal Matching	10	0.01	151	0.74 (0.38 – 1.46)
Variable Optimal Matching	10	0.001	89	0.83 (0.36 – 1.92)

Aim 3 Results Summary. This analysis was unable to detect a significant overall treatment effect for this sample when considering the risk for later maltreatment for a dropout group compared to a treatment group. Further, there were a large number of families who enrolled prenatally for whom birth data was not available and therefore CA/N reports were not able to be assessed. Therefore the focus of analyses was on those families with at least one postpartum visit. Further, with such a rare outcome event, attempts to match and assess a treatment effect had limited power to detect a difference particularly with sample reduction strategies such as propensity score matching.

When the outcome was considered a time-to-event variable and the time at risk was considered, the results were the same. Propensity score techniques successfully reduced measurable bias in baseline covariates and made the two treatment groups more similar allow for a more adequate comparison and estimation of the treatment effect. Even after propensity score

methods were incorporated, there was still a null finding for the treatment effect. Sensitivity analyses confirmed these findings across different propensity score matching techniques.

Subgroup analysis provided evidence for a protective treatment effect for families with a prior CA/N report in that the risk for a later report was significantly lower for treatment compared to dropout families. There also appeared to be differing effects according to the residential location of the families. In general there were null or positive effects for urban and suburban populations but rural populations appeared to be at higher risk.

Chapter 5: Discussion

The prevention of child abuse and neglect and the promotion of healthy child development are primary areas of focus for early childhood home visiting programs. However, prior research has found disappointing results in preventing official reports of maltreatment in most home visiting studies. A consistent concern in the literature is low family engagement with services. These two issues are particularly troubling given the broad expansion of home visiting services facilitated by a policy shift via the Affordable Care Act. As programs go to scale around the country and community-based agencies serve a greater number of families than have ever been reached, research is needed to better understand how to successfully engage families and prevent child maltreatment. Overall, the study results confirmed the challenges that scores of prior studies have highlighted. There appeared to be a lower than expected level of engagement and retention considering the level of service that should be provided in order to impact the outcomes defined by the program model.

There were, however, distinct differences in outcomes for certain families that have policy and program implications. Those that engaged prenatally had a much lower risk of maltreatment reports despite the fact that their level of risk was still high. Those that engaged during the postpartum period had higher rates of maltreatment. A significant overall treatment impact for those initially engaged with postpartum visits could not be detected even after applying advanced statistical techniques. While prior behavior might be considered more indicative of future behavior, a significant positive treatment effect was found for children with prior child welfare involvement. Results also indicated that preventing maltreatment appeared to be more challenging in rural areas. The following is a discussion of the findings for each research aim placing the results in the context of the current literature for that topic area. The

strengths and limitations of the study are provided followed by a discussion of the implications of for future research, policy, and practice.

Aim 1 Discussion: Service Utilization

The purpose of this first aim was to describe the service population. The families are organized into service populations by the agency across four groups, high-risk mothers, high-risk infants, teen mothers, and clients whose primary risk factor is poverty. By examining demographic characteristics and risk factors across these four groups it was apparent that the service populations do have some unique characteristics. However, it was also clear that there was substantial overlap on most of the risk factors across groups. It became apparent that a combination of individual risk factors or a cumulative risk score was more predictive of outcomes than program assignment. However, it may be clearer for funders to group families according to particular salient aspects of the child or caregiver. This then has implications for evaluation and research. If programs are created for ease of explanation or obtaining funds but these designations do not accurately reflect risk factors related to outcomes, it becomes crucial that data is collected that does map onto intervention components and outcomes outside these designations.

A cumulative risk index in prevention research is increasingly used to model the complex interactions of distinct risk factors. Studies utilizing this measurement approach have found that family risk factors related to caregiving and child development are inter-related and mutually interchangeable. Put most succinctly by Appleyard, Egeland, van Bulmen, and Sroufe (2005, p. 235), "the accumulation of risk factors, independent of the presence or absence of particular risk factors, impacts developmental outcomes, such that the greater the number of risk factors, the greater the prevalence of clinical problems." McRae and Barth (2008) found that a cumulative

risk score was able to identify children in need of mental health services with better specificity and sensitivity than case worker assessments. The cumulative risk score in this study was predictive of engagement, retention, and risk for maltreatment report in models controlling for demographic characteristics and service variables in a consistently strong linear manner. This indicates the utility of a cumulative risk score as predictive measure at the outset of services.

Further analysis could be conducted to identify the specific sets of risk factors and levels of risk that create the most sensitive and specific risk assessment tool for the agency's service population given the available data. For example, the section discussing aim three describes the use of a risk assessment scale based on epidemiologic studies linking birth records and CA/N reports. This scale is only based on only five risk factors and was able to categorize families with great prediction of risk for child abuse and neglect (CA/N) report in this study. Healthy Families America (HFA) uses the standardized Family Stress Checklist (Korfmacher, 2000) to assess risk. A scale similar to this one may be appropriate for NFN as well.

NFN service utilization. Consistent with other published home visiting research, this study found what could be characterized as relatively low levels of overall engagement with services. This issue is a major challenge for the field and research examining predictors of engagement can help identify barriers and inform engagement strategies (Ammerman, et al., 2006). A recent review (Damashek, Doughty, Ware, & Silovsky, 2011) described early attrition rates ranging from 20-67% as an issue that has "plagued" home-based maltreatment prevention. A meta-analysis (Gomby, 1999) reported 20-40% of families leave services before completion. Another review found 8-20% of service samples actively refused services and another 12-22% passively refused services (McCurdy, et al., 2006). Based on nurse report of termination reason,

it would appear that most cases closed due to what some would identify as passive refusal (missing appointments, not taking calls, moving without notifying of new contact information).

Of course, understanding issues of attrition in the present study was complicated by the fact NFN is guided by a case management philosophy rather than a set number of visits based on a set curriculum. Other models have very specific visit requirements that make assessment of adequate service participation fairly straightforward. In HFA for example, a service threshold of two years is set based on the program model. The NFP program model has a prescriptive schedule of visits from the early prenatal period through the child's second birthday. Understanding program dosage was further complicated by funding restrictions as there may be an approved number of visits depending on the client's referral source and provider. While NFN seeks to provide visits as needed regardless of payer source, the role of reimbursement for home visitation is an important consideration.

The research on how much is enough to produce results is unclear even for established, more structured programs. The NFP Elmira and Memphis studies, probably the closest program model to NFN, found that nurses completed an average of 9 and 7 prenatal visits and an average of 23 and 26 visits from birth to age two (Olds, et al., 1999). This represents about 50% of the 62 total visits offered by the program model (Kitzman, et al., 2010). In comparison, at the end of the two year program period, 38% of families in the NFP Denver trial had discontinued services (Olds, et al., 2002). In a review of 17 HFA service sites, the average family was enrolled in services for 14.8 months, received 31.2 visits, and 33% reached the service threshold of two years (Daro, McCurdy, Falconnier, & Stojanovic, 2003). Still, even given the level of typical attrition in these programs, the vast majority of NFN clients received less than half of the services provided in the other models in published studies. It should also be noted though that

these “evidence-based” programs are able to achieve positive results in outcome domains such as family economic self-sufficiency and maternal health with program participation much lower than what is prescribed by the model.

A final dilemma in assessing program utilization is the issue of time in the program and actual contact hours. About 20% were involved in services for more than 16 weeks. However, over 75% of families received less than ten hours of nurse contact. So, some families spread the intervention out over a longer period of time, but most families received a similar level of contact with the nurse. This may indicate underutilization in some cases but some families may benefit from knowing they have the option for contact over a longer period of time but may not want or need regular home visits. If a family still feels connected with the agency, they may be comfortable calling on the nurse for help as new concerns or problems emerge. The main leverage point of the intervention is the relationship between the nurse and the caregiver, which “is a process that gains depth and strength over time” (McNaughton, 2008, p. 407). In the present study, more visits did not necessarily predict better results, so certainly a better understanding of the relationship between service level, family risk, and nurse-client relationship is needed.

Predictors of engagement and retention. There were several variables that were significant predictors of both initial engagement and the number of visits a family received. Prior research has indicated that a behavior such as engaging in a voluntary home visiting program is predicted largely by the individual’s intent to do so (Fishbein, et al., 1997; McCurdy, et al., 2006). While this seems straightforward, it helps to examine barriers to engagement by exploring what factors would impact an individual’s intent to receive home visits. Overall, these findings on participation were consistent with the Theory of Parental Involvement (McCurdy &

Daro, 2001). This theory predicts that there are levels of influence across parental, provider, and program levels of ecology with the most potent being the individual cost-benefit calculation.

This study found that while the bulk of the variance in engagement that could be explained was at the child and caregiver level, the nurse that is assigned to a given participant exerts some influence on their level of service use. Research on mental health services have found that non-specific treatment effects, including the client therapist relationship may have equal or even more importance than specific program components (Messer & Wampold, 2002). These "process variables" may have more of an impact on outcomes than individual factors prior to treatment (Kolb, et al., 1985). One study found that a structured assessment examining the way that the Parent-Child Interaction Therapy therapist talked to the caregiver in the first thirty minutes predicted dropout versus completion (Harwood & Eyberg, 2004). A number of recent studies have tried to unpack this idea of a therapeutic alliance and increasing attention has been paid to this issue in home visitation (Korfmacher, et al., 2007). More research is needed to include nurse-level variables such as race/ethnic match, experience, and measures of quality of relationships with the clients into these models.

There were factors within the caregiver level that predicted initial and ongoing engagement. While it may be intuitive to consider higher risk families harder to reach, maternal history of child abuse and neglect was a consistently positive indicator of greater engagement and retention. This finding may be interpreted as greater receptivity to parenting support for caregivers with a history of personal trauma. While not a predictor of initial engagement, mothers who began services and reported current domestic violence received a greater number of visits. It is not clear if this reflected a greater motivation on the part of the caregiver due to perceived protective presence of the home visitor or greater efforts on the part of the nurse due to

concern for the caregiver and infant's safety. Caregivers who had previously experienced a fetal death and those with low birthweight or premature babies were more likely to engage with services. This may reflect heightened concern for the health of the baby and perceived benefit of having a health care provider in the home.

The findings of this study suggest that the desire to receive home visits is likely linked to a perceived level of risk or the realized potential of a possible untoward outcome related to the current pregnancy. Mothers initiating services prenatally with higher levels of caregiving stress received significantly more visits. Those who had a prior pregnancy complication were more likely to engage with services. This likely reflects a short-term cost-benefit analysis by the caregiver based on the assumption that this service is worth the time investment, assuming that it will decrease the chance of an identified negative outcome or the visits will help buffer the stress. The present findings are consistent with other research that characterizes increased engagement for mothers with a low birth weight child as an example of the mother's awareness and responsiveness to their infant's needs (McGuigan, et al., 2003; McCurdy, et al., 2006).

So if greater risk is associated with engagement the question becomes, "What is associated with refusal or avoidance?" There are two common positions that explain the "avoidance" of home visiting that deserve further discussion. As outlined by McCurdy, et al., 2006, one explanation suggests that a family avoids services, or does not engage, because they have accurately assessed themselves as having adequate parenting skill and knowledge to handle their caregiving demands. This is the disengagement as "informed consumer choice" explanation (McCurdy, et al., p. 1196). The other explanation suggests that a family who avoids the home visitor is also the one who is more likely to be isolated and at higher risk for outcomes such as child abuse and neglect that the intervention is attempting to address. The first theory

would suggest that the 20% of the caregivers who decided to drop out after one visit were making an informed decision that was superior to that of a trained professional in most cases.

While qualitative data on the caregiver perspective was not available for the present study, there is perhaps more evidence for the informed consumer idea of engagement. All of the families referred to this program face some barrier or risk. Families with higher levels of risk were more likely to engage in services. The more geographically isolated families in rural areas are actually more likely to receive visits in the postpartum referred group. This may suggest that higher risk families are able to accurately recognize the potential benefit of additional support for parenting. On the other hand, families with an infant with a prior CA/N report had lower odds of engagement and fewer overall visits. Thus, it may be not only be the level of risk for maltreatment but the level of formal involvement with other systems or concern with service providers that impacts engagement. For families with a lot of formal system involvement, there may be a negative perception of increased surveillance. Whereas for families with no formal system involvement, the potential perceived benefit may not be impacted by perceived risks of participation. Further, there was not a measure capturing the caregiver's motivation for services. This construct would likely be important in the decision to engage in services and important to include in the propensity score model.

Dose of service. Other home visiting studies have attempted to explore dose as a moderator of outcomes. Studies of NFP (Korfmacher, et. al., 1998; Olds, et al., 1990) and Parents as Teachers (Wagner & Clayton, 1999) have found that engagement is positively related to program outcomes. In contrast, the Healthy Families Alaska evaluation (Duggan, et al., 2007) found no evidence of impact on child abuse outcomes when comparing families with a high-dose of services to the control group. However, virtually all studies to date, including the current one,

have been correlational without randomizing families to different levels of expected involvement (Korfmacher, et al., 2007). Propensity analysis has been suggested as one way to rule out threats to validity stemming from client characteristics given the difficulty in randomizing families to different levels of service (Korfmacher, et al., 2007).

In this study, several barriers to the measurement of dosage were uncovered. The analysis investigating the relationship between number of visits and skill delivery found a significant but potentially weak relationship riddled with confounding explanations. Nurses delivered a much higher number of skills during the first visit than expected. Depending on the family, the nurse report of skills delivered might be the same for a dropout and a treatment family. What families receive in addition to these skills appears to be dependent on the level of need and level of risk of the family. Since the number of visits is driven by the nurse and family, and was not randomly distributed, it is difficult to untangle the relationship between dose of service (whether measured as visits or skills provided) and later outcomes. Overall, these findings should be considered inconclusive when examining the effect of dose of service on later outcomes.

Aim 2 Discussions: Service Subgroups

An area of inquiry in home visiting research that has received relatively little emphasis is concerned with characterizing different risk groups and then determining for whom the intervention was the most effective (Olds & Korfmacher, 1998). This aim was concerned with examining rates of child maltreatment, child developmental concerns, and maternal stress and depression across different participant subgroups that are considered to place them at higher risk for adverse outcomes. There were specific family subgroups that did have a statistically significant effect for improved child maltreatment outcomes and other groups that showed

greater risk for adverse child and maternal risks. The results of effects in multivariate models predicting CA/N reports and specific subgroup analyses will be discussed in this section by key subgroups of interest.

First visit timing. Families who had at least one prenatal visit were much less likely to have a later maltreatment compared to those who only had postpartum visits. This was the largest effect found in this study. The effect size of the relationship predicting a later CA/N report comparing those with no prenatal visits and those with at least one was OR = 20.4 (95% CI 10.1 to 41.3). Only 1% of the 712 prenatally engaged families had a later report compared to 18% of the 2,664 families who began visits postpartum. Unfortunately it was not possible to assess whether prenatal-only home visitation was more or less effective than prenatal plus postpartum intervention, due to the lack of identifying information for families who dropped out prior to the birth of the baby.

In attempting to explain these results, the demographic characteristics of families who receive prenatal visits were compared to those who initiated in the postpartum period. The main difference between these groups was referral source. Those with a prenatal visit were more likely to be self or relative referral (40%) compared to only 10% that were self-referred postpartum. Families who initiate service prenatally are likely seeking out services for themselves and may have some existing knowledge about the agency. These mothers can be thought of as a purely preventive group on one hand as they are more likely to be looking for services because they are proactively concerned about their future and the health of their child. On the other hand, there were significant pre-existing risk factors within this self-referral group compared to the post-partum initiators. For example, these mothers were more likely to be homeless, be currently experiencing partner violence, have a personal history of child abuse or

neglect, have a prior fetal death, and have less than 18 months since their last pregnancy. For these families, it appears that the mother has made an assessment of her personal level of risk and has chosen to initiate services to address these risks.

Those who initiate services during the postpartum, which are those most likely to be referred from the hospital setting, are more likely have a child in the NICU or special care nursery, are more likely to have a small or early baby, and more likely to have a drug exposed baby. At this point, services are in many ways a reactive response to the needs of a high-risk baby initiated by a health professional responding to a perceived risk. Further, both substance abuse and maltreatment history have been associated with higher risk of early childhood maltreatment (Appleyard, et al., 2011) and infant drug-exposure may well be a proxy for a substance abuse problem. The mother may or may not actually be interested in services to address any of her own caregiving risk factors but is concerned about the well-being of the child.

Prior report recurrence prevention. This group was not listed as an a priori subgroup for analysis because it was assumed that there would be too few children with CA/N reports between birth and initiation of home visiting services. However, 2.2% of the sample had a CA/N report then began home visiting services. Of course, it should be noted that measurement of CA/N was limited to the infant-level of prior report. It is possible that a larger percentage of caregivers had a prior report on an older child. Although nurse documentation was inconsistent, another 7% of caregivers had some form of prior DFS contact in their risk factor documentation. In this context, services are no longer considered primary prevention but are framed as the prevention of recurrence of maltreatment. Prior report of maltreatment is one of the most consistent predictors of future maltreatment and represents a group of families for whom services are greatly needed (DePanfilis & Zuravin, 1998; DePanfilis & Zuravin, 2002). Using NCANDS

data, Palusci (2011) found a recurrence rate of 9.2% for children aged 0-4, with about one-fourth of the recurrence for infants occurring within six months of the first report.

The current findings are noteworthy given that prior studies have found that recurrence is most common among younger children and families with higher level of risk (Berrick, Needell, Barth, & Jonson-Reid, 1998; Drake et al., 2002; Drake, Jonson-Reid, & Sapokaite, 2006). In a sample of low-income children in the same geographic area, Drake, Jonson-Reid, & Sapokaite (2006) found a three year recurrence rate of 49.9% for children ages 0-1 at the time of the initial report. This is an almost identical rate to what was found for the one visit dropout group in the current study. Given the lack of a true comparison group, these results serve as an adequate comparison for the expected re-report rate in the time frame studied for this sample. While more visits was associated with higher risk of report for most families in the study, higher dosage of visits had a protective effect for those with prior CA/N reports.

There is some debate in the home visiting literature whether or not to focus on intervening with families who have already maltreated a child or focus specifically on primary prevention. Olds, Eckenrode, & Kitzman (2005) have formally stated that the goal of NFP is to prevent initiation and not recurrence of maltreatment and have targeted first-time mothers for their intervention for whom it was highly unlikely that they had a previous report, at least for not for their own children. Therefore it is not possible to examine prevention of recurrence with the NFP model.

A Canadian home-visiting study of public health nurses (MacMillan, et al., 2005) randomized families with a child under age 13 recruited from the children protection agencies with a recent history of abuse or neglect. There was no significant difference between treatment and control groups with about half of the sample returning for a neglect report and a third for

physical abuse over a three year follow-up. In fact, based on hospital records, a higher percentage of the treatment group (24% vs. 11%) had hospital stays adjudicated as physical abuse or neglect. The study also found no difference on several observational scales of child abuse risk.

A randomized study of Healthy Families New York (HFNY; Dumont, et al., 2008) identified 20% of their sample as having a prior child abuse or neglect report. In the final HFNY report (Dumont, et al., 2010), a Recurrence Reduction Opportunity (RRO) subgroup included these women who had a confirmed report within the past five years. In a seven year follow-up, RRO mothers had lower rates of confirmed CPS reports compared to the control group (42% vs. 60%). This finding was not quite significant at $p < .10$, but a relatively low cell size for the total RRO (n=104) group diminished the power to detect a difference. There was a significant difference for the rate of child welfare preventive services that were initiated for the HFNY group compared to the control (38% vs. 60%). Given the consistency of the findings of the current study and those with the HFNY analysis of maltreatment recurrence, there may be an opportunity to think about focusing home visiting services not just for the primary prevention of maltreatment but also intervening with families who have already come to the attention of the system.

Maternal parity. The dominant nurse home-visiting model, Nurse Family Partnership, exclusively serves first-time mothers on the assumption that these are both the most unprepared mothers, yet least defensive, and that the intervention will translate to future children. However, the model has never been tested with multiparous mothers and has contributed to the assumption that first-time mothers benefit the most from home visitation services (Galano & Huntington, 2012). A recent meta-analysis found a significant effect of family size on risk for abuse and

neglect (Stith, et al., 2009), however the question of the differential effectiveness of home visiting for multiparous compared to primiparous mothers has been entirely lacking from most studies and meta-analyses examining maltreatment prevention (Galano & Huntington, 2012).

In this sample, mothers who already have children in the home were 50% more likely to have a later CA/N report compared to first time mothers. Multiparous mothers had higher levels of stress and were more likely to be in the clinical range for postpartum depression. Their children were also more likely to have a development concern measured by the ASQ. With over half of the sample experiencing a previous pregnancy, mothers with children represent a large at-risk group in the sample. They would unfortunately not be eligible for services provided by the NFP model.

One reason why parity may be related to risk for maltreatment is related to an increase in parenting stress associated with caring and providing for multiple children in the home (Chaffin, Kelleher, & Hollenberg, 1996; Shook Slack, et al., 2003) and the large effect between parenting stress and risk for maltreatment (Stith, et al., 2009). Post hoc analyses in this study yielded a dose response in the number of siblings in the home and the risk for later report. This is consistent with prior literature in the field. Bae, Solomon, & Gelles (2009) found that family size increased the risk of maltreatment recurrence, with each extra dependent increasing the risk for a later report by 16%. Population-based birth cohort studies confirm this finding. Putnam-Hornstein & Needell found a doubling of the risk (adjusted RR = 2.0) for later report with 3 or more children in the home compared to a singleton. Wu, et al., (2004) found a similar effect size (adjusted RR = 2.7).

The results comparing outcomes by parity in home visiting studies are scant. One study found a reduction in the EPDS scale for primiparous mothers but not for multiparous mothers

(Armstrong, et al., 1999). Using data from Healthy Families Virginia, Galano & Huntington (2012) provide the only study known to examine the question of parity. This study found that multiparous mothers were older, more likely to be African-American, more likely to be employed full-time, had higher overall risk scores, and were better educated. Parity was not related to program participation or outcomes including immunization, HOME scale scores, or subsequent birth rates. The authors conclude that there is no evidence that mothers with children respond less favorably than first-time mothers, nor are they any less at risk. In a follow-up qualitative portion of the study, evaluators asked program managers about specific challenges working with multiparous mothers. The quote that emerged in the findings and was often repeated was, "All of the things that are challenging and stressful about parenting are made more so by having multiple children." (Galano & Huntington, 2012, p. 68). There is a clear need to consider the targeting of services to first-time mothers at the exclusion of those who may be in most need.

Newborn health status. This study found higher rates of child maltreatment and developmental concerns among children who were low birth weight, premature, or medically-fragile. However, there was not a difference in levels of caregiver stress, postpartum depression, or self-reported mental health issues. The findings for higher rates of maltreatment and developmental concerns among preterm and low birth weight newborns is consistent with previous literature (Ashdown-Lambert, 2005; Olivieri, et al., 2012; Parrish, et al., 2011; Putnam-Hornstein & Needell, 2011, Salt & Redshaw, 2005, Wu, et al., 2004), while the similar rates of depression and stress deviate from prior findings.

Previous research has found a marked increase in the level of psychological distress, particularly parenting stress, for caregivers of children who are preterm, low-birth weight, or

have associated medical risks (Singer, et al., 1999). Parents of low birthweight infants often experience posttraumatic stress disorder (Jotzo & Poets, 2005), with no significant reduction in symptoms more than a year after birth (Kersting, et al., 2004). A follow-up study of low birth weight infants at age eight found persistent differences in maternal stress and a range of family outcomes (Singer, et al., 2007). The emotional impact of caring for a fragile child likely contributes to an increased risk for child maltreatment. Among families with low birthweight children, research suggests that risk for maltreatment is driven by parental risk factors such as anxiety symptoms and not perinatal risk factors such as physical disability (Strathearn, Gray, & Wood, 2001; Zelkowitz, Bardin, & Papageorgiou, 2007).

The lack of relationship between medically-fragile, low-birth weight, or preterm infants and caregiver mental health was not expected for this study. It is possible though that referral to NFN services made an impact on these outcomes to the degree that they were decreased at the time of screening. Without an adequate baseline comparison group, it is difficult to determine to what extent the families with medically fragile infants entering services are similar to those that do not engage with NFN.

Maternal behavioral health. This study found an increased risk for CA/N report, gross motor skill developmental concerns, postpartum depression, and caregiving stress among caregivers with mental health and substance abuse concerns. In most cases, the risk for these outcomes was increased for those evidencing co-occurring disorders. Behavioral health outcomes were determined by maternal self-report and were likely underestimated in this study.

An increased risk for maltreatment and potential indication of child development concerns are consistent with prior literature exploring the impact of maternal mental health on attachment and caregiving. Maternal mental health status has been linked from theory and with

empirical evidence with higher risk for child abuse or neglect during early childhood (Belsky, 1984; Howard & Brooks-Gunn, 2009; Ammerman, et al., 2011; Chaffin, Kelleher, & Hollenberg, 1996; Windham, et al., 2004; Conron, et al., 2009; Casanueva, et al., 2011). Exposure to prenatal maternal depression and stress can alter the development of crucial fetal neuroendocrine systems while postpartum symptoms can impact maternal-child interactions and disrupt healthy attachment (Hammen, Shih, & Brennan, 2004). While maternal substance use, depression, anxiety, and stress can all be individually linked to poor family outcomes, the reality is that mothers often experience symptoms simultaneously, compounding risk (Ammerman, Putnam, Chard, Stevens, Van Ginkel, 2011).

Aim 3 Discussion: Maltreatment Prevention Effectiveness

Through the use of propensity score (PS) matching techniques, this study was able to decrease a large portion of the bias that existed in selected covariates between the treatment and dropout comparison group as defined in this study as those receiving only one visit. A limitation of this effort was that the absence of child level information on prenatal only service users limited this analysis to those who had at least one postpartum visit. When viewed as a whole, no significant treatment effect was found after PS matching. Further, a dose effect of increased visits could not be detected for families who received additional home visits after controlling for child and family risk factors that could be driving higher service utilization. These findings were consistent when examining only families who began services during the postpartum period.

The null findings for maltreatment prevention from this study should be placed in the larger context of the current literature. The most recent meta-analysis of published home visiting research (Filene, 2012) built on three prior meta-analyses (Bilukha et al., 2005; Gomby, 2005; Sweet & Appelbaum, 2004) by exploring both program outcomes as well as components of

different home visiting models that were associated with better outcomes, including child maltreatment prevention. This meta-analysis included eight studies that examined child maltreatment as a post-test outcome and calculated a non-significant effect size of .02 (95% CI -.08 to .11) across these studies. There were a total of 23 studies that included child maltreatment as any outcome at any time point in the study and again the effect size (.04, 95% CI -.01 to .09) was non-significant.

Despite the non-significant overall effect size, the Filene (2012) meta-analysis examined the specific program components across studies that were associated with improved child maltreatment outcomes. The largest effect sizes for components associated with successful programs were those that included discipline and behavior management, stress or anger management, and promotion of children's language development. Delivery methods that were associated with better outcomes were programs that used a standardized curriculum and did not have professional home visitors. In contrast, a previous systematic review of all home visitation programs reported a reduction of reported maltreatment by 39% with visitation by nurses or mental health workers yielding greater effects than by paraprofessionals (Bilukha, et al., 2005). The match between visitor and client on race/ethnicity and offering program delivery in a language other than English were associated with larger program effects for child maltreatment outcomes.

Nurses for Newborns services is distinct from some of the more structured programs and tends to operate as a short-term crisis and case management program for most of the very high-risk caregivers and medically-fragile newborns. A great deal of what the nurses are accomplishing with the family is related to the physical health of the newborn and the needs of the mother to support their health. These efforts include helping the mother secure a medical

home, insuring proper feeding and health care practices, and monitoring the overall health of the infant. While initial care is certainly related to child abuse and neglect prevention in the short-term, maltreatment prevention in the long run is likely associated with additional caregiver factors that are not a direct focus of the program. A key finding from the meta-analysis was that greater effects were found for those programs with a standardized curriculum that include components relating to child development or parenting beyond infancy. NFN does not use a standardized curriculum and the nurses have a great deal of autonomy to provide services to their clients within broad clinical guidelines.

Although there are a set of teaching skills that form the clinical protocol, the likelihood of coverage of these skills varies. According to nurse documentation, over 90% of families served receive teaching skills on "child abuse and neglect prevention" while only 29% receive the shaken baby syndrome teaching and 66% receive the growth and development teaching skill. There do not appear to be any specific teaching skills around the domains identified by the aforementioned meta-analysis including discipline or behavior management, stress or anger management, or promotion of child language development. This is related to the fact that these areas are not likely to be related to the immediate needs of the family. Unless there are older children in the home, the nurse is probably not addressing issues related to discipline or language development with a newborn.

Stress management is addressed in the clinical guidelines by psychoeducational approach that includes informing the client about the impacts of stress on health and the health of the fetus. Materials related to stress reduction techniques are provided to the mother. Other areas of the caregiver's risks related to stress are likely also addressed by teaching skills around building a support system (85%), rest (64%), exploring educational options (47%), or referrals for job

training (23%). However, there do not appear to be any specific evidence-based anger or stress management interventions. There is a teaching skill around crying tips (64%) and in the clinical guidelines nurses are instructed to provide the Prevent Child Abuse handout "Twelve Alternatives to Lashing Out at Your Child" and make appropriate referrals including Children's Division reports. However, there is no behavioral component that seeks to improve parent-child interaction or improve disciplinary strategies (which become more important as the child enters the toddler and preschool years) such as those found in models such as PCIT or Triple P parenting programs.

As mentioned earlier in the discussion, it is evident that outcomes do vary by nurse assigned. There was no ability to measure non-specific treatment effects. The agency hires nurses by the service needs in a given zip code. So there is likely to be a good ethnic/racial match, but not necessarily. The findings for Hispanic clients who receive the services of a community health educator indicate some preliminary evidence that providing culturally-sensitive home visitation that addresses specific needs relating to language or other cultural barriers is likely to improve program engagement and may improve later outcomes. This is an area of the service model and home visiting services that should be explored further in future intervention research with culturally diverse clients.

Population-based Maltreatment Rates. One limitation of this study is the lack of a randomized, controlled comparison group to determine whether or not maltreatment rates for families receiving NFN is higher or lower than what would be expected in the absence of services. Overall, about 15% of the infants in this sample had a later report. Is this rate high or low? The study attempted to use a dropout comparison group among families who only had one home visit. However, there are limitations to this approach. Services in this geographic are

well-saturated, most high-risk mothers and their babies in the area are at least referred to this specific agency. This makes identifying a similar community-based comparison group difficult as well.

There have recently been a set of studies that have utilized population-based birth data to determine rates of CA/N using risk factors from birth records and linked child welfare administrative data. While the reporting system and populations of interest are different, these data provide one way to determine whether or not rates of CA/N in this sample are different from what can be found in a general population. There is no way to know whether the families in these studies also received home visiting services, but this provides a reasonable marker of what can be expected for families with similar levels of risk.

Putnam-Hornstein & Needell (2011) followed a birth cohort in California for five years linking vital birth records with child protective service (CPS) contact and Wu, et al. (2006) followed newborns in Florida for one year merging birth records and CPS reports. As a post-hoc analysis, the rates of CA/N across three risk groups were compared to these two studies. For comparison with the Wu study, the rates of CA/N prior to age one (n=3,373) were computed for families in the sample that were low birth weight, had greater than two siblings, had no high school education, and had inadequate prenatal care. The five year follow-up period for the Putnam-Hornstein study is longer than the follow-up period families in this sample, so this comparison included families that could be tracked for reports through age three (n=1,575). This analysis uses a slightly different framing of the time-to-event analysis. In prior analyses the time to event was tracked from termination of services. For this analysis, the time is tracked from the child's birth date through the first or third year of life. Also, previous analyses considered

whether or not the first report occurred following services. This analysis examines any report before or after services were initiated.

Before comparing the results of the rates of CA/N between these birth cohorts, it is important to examine the relative level of risk in a general population birth cohort and the study cohort of families receiving home visiting services that can be tracked from birth. The following pie charts in Figure 17 compare the distribution of risk factors based on the Wu, et al., (2004) risk profiles. This epidemiologic risk assessment tool was constructed using five factors (mother smoked during pregnancy, more than two siblings, Medicaid beneficiary, unmarried marital status, and infant born low birth weight). Similar measures are available in the current dataset, although drawn from self-report and not from administrative birth records. Families with 4-5 factors are extremely high risk, those with 3 are high risk, those with 2 are average risk, and those with 0-1 are low risk.

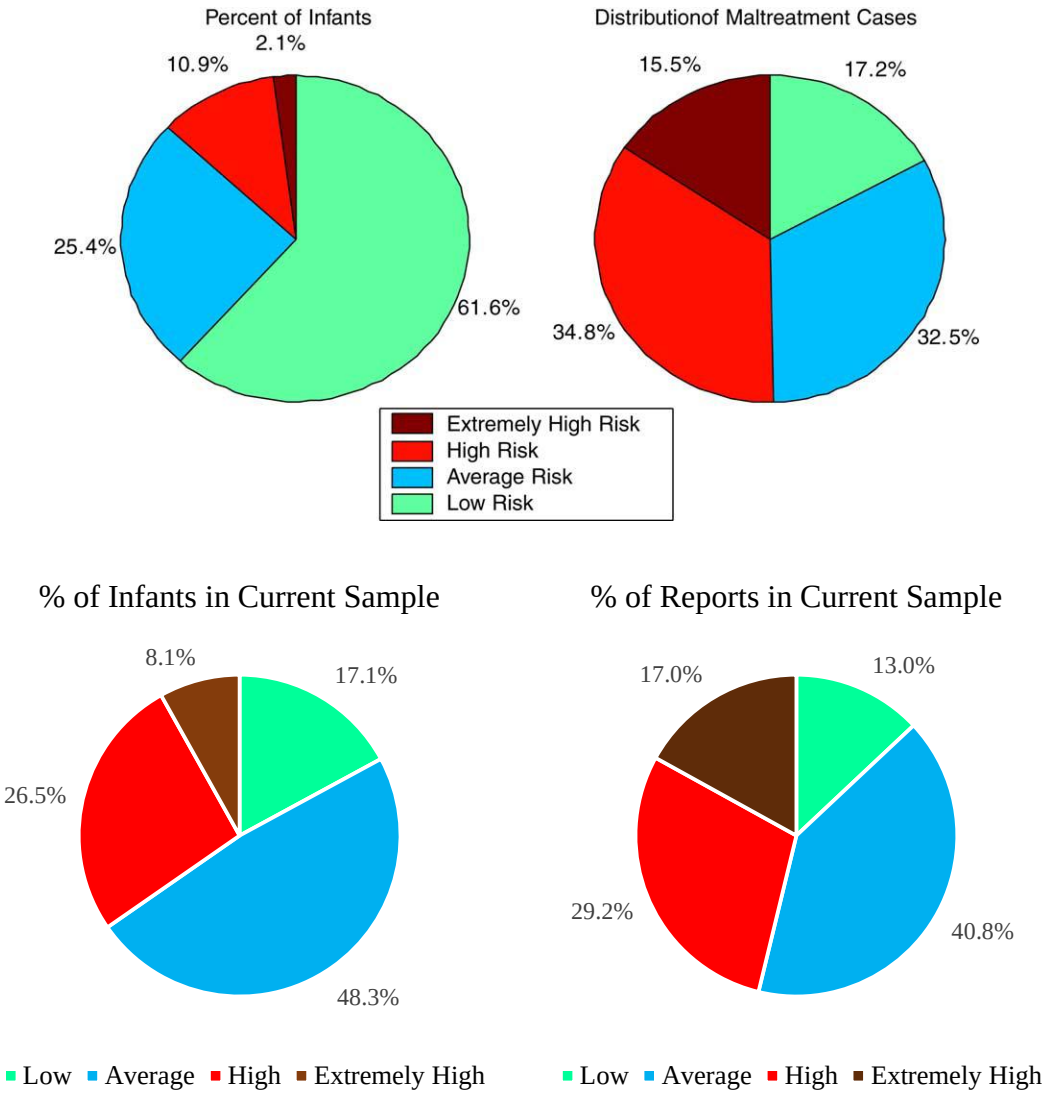


Figure 17. Percent of infants (ages 0-1) across four risk category groups compared to the distribution among reported cases of child maltreatment. The top two pie graphs are from Wu, et al, 2004, Figure 1 cohort of 189,055 Florida births in 1996 and the bottom pie graphs represent the current study of a home visiting service population (n=3376).

The first clear difference between the birth cohort and the current sample is the difference in the risk profile shown in the left pie graph. Well over half of the families in the Florida birth cohort are considered low risk compared to 17% of the present study sample. Over 65% of the current sample fell into the high-risk category compared to 13% in the birth cohort study. It is

clear that the NFN sample represents a much higher risk group of families than what is likely found in the general population. However, the distribution of maltreatment cases across the four risk groups shows a striking similarity between the two samples. Although these samples are drawn from different points in time in different states, they are able to predict with remarkable accuracy the risk composition of the reported cases of maltreatment.

With the differences in sample composition in mind, the following table provides a comparison of the rates of child welfare contact at one year compared to the Wu and colleagues findings and the rate through age three to compare with the Putnam-Hornstein & Needell findings through age five. This table compares the report rates based on Medicaid status, infant low birth weight, more than two siblings, prenatal care, race, and risk status. The overall rate of CA/N is higher at both ages compared to the two cohort studies. Again, this is a comparison between the general population and a high-risk service population. The Wu, et al., 2004 only reported total reports for all types for "verified" cases only, so for this comparison only "substantiated" cases were counted. The Putnam-Hornstein & Needell study uses all reports regardless of disposition, so all reports were used for this comparison.

While the rate of maltreatment from a general population birth cohort is expected to be much lower than the current sample given the high-risk nature of the families served, the rate for Medicaid birth payment is almost identical to the rate for the current sample with infant Medicaid. Given that about 85% of the current sample has infant Medicaid, this appears to be a useful figure for comparison. At year one, the rate was 1.5% compared to 1.6% and the age three Medicaid rate was 20.1% compared to 21.2% at age five.

Table 25
Comparison of Risk Factors and Rates of Reports at Age One and Age Three for the Current Study and Two Population-Based Epidemiologic Studies

	Wu, et al., (2004)	Current Study Substantiated	Putnam-Hornstein & Needell (2011)	Current Study Any Report
Reports through Age	1 year	1 year	5 years	3 years
Overall Rate	0.9	1.4	13.9	18.9
Medicaid Payment	1.5	1.6	21.2	20.1
Infant LBW	2.1	2.1	18.7	22
>2 Siblings	2.2	2.3	20.7	27.1
No High School	1.9	1.6	20.1	24.3
No Prenatal Care	-	-	48.9	40
Inadequate Prenatal Care	2.2	1.5	-	-
Race White	0.7	1.9	13.3	24.2
Race Black	1.4	0.9	30	12.7

Although some findings are almost twenty years old at this point, another useful comparison for this group is the NFP trials that include only mothers enrolled prenatally continuing visits through the postpartum period. In the Elmira trial, there was not a significant treatment effect for the entire sample. However, 10% of the comparison group and 5% of the treatment group had verified ("indicated") CA/N reports two years after follow-up. The rate of substantiated child abuse and neglect based on "pretest and pilot work" was reported to be 3-4% for low-income firstborn children in Memphis and was considered "too low to serve as a viable outcome measure" (Olds, et al., 1999, p. 57). Similarly, the "low rates of state verified cases of child abuse and neglect" made it "impossible" to look at child maltreatment outcomes in the Denver NFP trial (Olds, et al., 2002, p. 488). Thus, the current state of the literature does not provide an adequate comparison figure for what the expected rate of maltreatment would be for families initiating services prenatally.

The most recent study examining maltreatment rates for this type of population is a from a statewide randomized effectiveness study of Healthy Families Massachusetts (Easterbrooks, et al., 2013). This model provided paraprofessional home visiting to first-time teenage mothers.

After a two year follow-up, 29% of all families in both conditions had a later report. Similar to the current study, this evaluation found that maternal history of maltreatment victimization and current intimate partner violence was predictive of later reports at the bivariate level. This study found that families in the home visiting treatment condition were more likely (OR = 1.72) to have a substantiated report of maltreatment compared to the referrals and information only group. The authors attribute this finding to increased surveillance, a potential limitation addressed in the next section.

Strengths and Limitations

While the main limitation of the study is the lack of control over design and data, this lack of investigator manipulation is also one of the strengths of the project. Instead of examining what works under optimal conditions, this type of community-based study focuses more on external validity by examining whether or not interventions are "palatable, feasible, durable, affordable, and sustainable in real-world settings" (Jensen, Hoagwood, & Trickett, 1999, p. 207). The practicality of this study is enhanced by the agency's pre-existing case record system rather than reliance on original data collection. The system is unique in that it not only includes quantitative information like number of visits or maternal depression scores, but also includes open-ended clinical nursing notes.

There are several limitations of this dissertation that should be noted. Randomized experiments are the gold standard for examining the causal effects of a treatment. Thus, a general limitation of the study is the quasi-experimental design in assessing a treatment effect. This limitation is at least partially overcome by the fact that the evaluation examined an implementation of a home-visiting program as it currently exists deployed by a community agency without research manipulation. While this may lead to biased estimation of precise

treatment effects by introducing threats to validity, this type of research is currently needed to move policy and practice forward in the area of maltreatment prevention. The literature has benefited from several RCTs of home visiting services delivered under highly controlled conditions. The families served by NFN represent a much higher and diverse risk pool than what has been examined in these studies. These families likely reflect the types of families that will be served as home visiting program are taken to scale. However, given the demographics of the St. Louis area, results cannot be generalized to racial or ethnic subgroups beyond White or African-American.

Measuring maltreatment. Another limitation of the study is the method of measuring the main outcome of interest, the prevention of abusive or neglectful caregiving behaviors. This study utilized official reports to child welfare services to indicate the presence of abusive or neglectful parenting. Similarly, the lack of a report was an indication of an absence of maltreatment. Some have questioned the reliance on official reports of maltreatment and administrative records to measure abuse and neglect (Olds, 2005). There is reasonable concern that this may lead to a significant underestimation of true cases of maltreatment that fall below a subjective threshold to warrant suspicion. Further, the administrative records match was made at the child level. It is quite possible that the caregiver may have had another CPS report for another child during the study period.

On the other hand, the use of official reports has some strengths. Some researchers have used retrospective reporting of child maltreatment as a means of identifying a sample that has experienced abuse and neglect. However, adult reporting of maltreatment that may or may not have occurred during childhood is fraught with issues of reliability and validity. Differences in outcomes have been found in studies depending on whether official reports or self-reports of

maltreatment are used (Widom, et al., 1999; Widom, Raphael, & DuMont, 2004). Child abuse potential scales may not adequately capture behaviors and questions are geared to parents of older children. Observational measures are more expensive and may miss behaviors due to bias introduced by being observed.

Surveillance bias. An often-cited concern in studies examining official reports of maltreatment in home-visiting programs is the issue of surveillance bias (Olds, et al., 1993; Duggan, et al., 2007; Biluhka, et al., 2004; Barlow, Simkiss, & Stewart-Brown, 2006). Specifically, the presence of a home visitor in the home of a subject in the intervention condition increases the likelihood that child maltreatment will be observed and thus reported. Two studies examined the number of reports by the home visitor compared to the reports of both the visitor and CPS. The results of these studies indicate that bias increases the rate of reported maltreatment by 80% (Brayden, et al., 1993) and 150% (Dawson, Van Doornick, & Robinson, 1989). When included in a meta-analysis, the results suggest that the presence of a visitor bias in combination with positive results of reduced maltreatment reporting only strengthens the conclusion by underestimating the true effect on maltreating behavior. Chaffin & Bard (2006) examined two outcome evaluations of interventions in child welfare population and concluded that surveillance had a small net effect and that home-visiting services were not more biased than center-based services. They also conclude that surveillance effect may be substantial when families are actively involved in services but are "washed out" given high attrition and a sufficient follow-up period (Chaffin & Bard, 2006, p. 309).

While delivering home visiting services, visitors are required to report suspected cases of child abuse and neglect. This might have two effects in this study. As suggested in prior research, the nurse may avoid certain questions to avoid a "chilling effect" from the threat of

reporting (Dawes-Knight, et al., 2006). A second concern is that families who have been reported to CPS might drop out of services. These are likely to be real concerns with nurses intervening with high-risk families. They are likely concerned that a CPS report will damage rapport despite well-meaning intentions for the safety and well-being of the child.

The general sense from discussion with agency members is that surveillance bias may actually operate in the opposite direction. Nurses are aware that they may be the last line of defense for a family and will, if anything, delay a CPS report in favor of attempting to work with the family. The data was consistent with this as less than 5% of families had a hotline call documented by the nurse and 4.2% had an official report in the CPS data during the time period between the first visit and case termination. This study was able to disentangle reports prior, during and after which decreases the risk of surveillance bias by the nurse. It is not clear whether a nurse serving a higher risk family may be more successful in engaging them in services outside NFN. If this is the case then surveillance bias might still be in play related to a different provider. However, in a study addressing the issue of surveillance bias, low-income children reported for CA/N appeared to be accurately identified as those with significant needs (Jonson-Reid et al, 2009).

Study Implications

This research has a number of benefits for policy and practice in the area maltreatment prevention and home visitation. The primary benefit is found in the partnership with an agency like NFN and the families they serve. While much work has been done in the study of home visitation under controlled conditions, NFN has been implementing services in less-than-ideal conditions for over two decades. NFN's laboratory is the real world. As home visitation and maltreatment prevention moves from the context of the theoretical to the applied, programs such

as NFN must be included in the scientific discourse. This study provides the opportunity to learn from an established program that struggles with the balance of organizational momentum and friction. For programs across the country grappling with the prospect of building and growing new programs, this study will provide a valuable model of what can be expected in the long-term.

NFN seeks to engage and help the families have been identified by professionals in the community or by the clients themselves as most in need. This often means working with caregivers facing extremely challenging circumstances. While other programs have shied away from mothers with mental illness, disability, and substance abuse problems or from infants with serious medical conditions, NFN has actively pursued them. These populations certainly do not make the best candidates to establish the efficacy of an intervention, but they cannot be ignored when planning to take to scale programs that seek to protect children and support high-risk families. This research provides an indication of what is possible in the field of home visitation and also the possible limits of applying this type of intervention.

This study also provides valuable information regarding the efficiency of home visiting services by exploring the outcomes associated with level of service driven almost entirely by the nurse and client's determination of need. There is still much unknown about how to provide minimally-sufficient services based on the needs of the family. Home visiting is not a one-size-fits-all program. While the importance of developing a strong model that can be delivered with fidelity is incredibly important, this cannot come at the expense of the flexibility and professional judgment required to best serve families. Research can provide better indications of what level of intervention seems to produce optimal results. The goal is then to maximize results

by providing the lowest level of intervention required. This equation will differ across types of families.

The overall goal of this research and the anticipated benefit is to contribute to the larger discussion of early childhood home visitation policy. At the time of this writing, the Title V Maternal & Child Health Services Block Grant received a $50 million cut to the original appropriations to support home-visiting services. The current economic downturn impacts government programs but also disproportionately impacts families at the lower end of the socioeconomic spectrum. The health and well-being of the infants and children are completely in the hands of their family and the government programs that provide needed support. Agencies such as NFN stand directly between families and policy, linking and delivering resources with those who need them. This study provides one way that the voice of the agency and the families they serve can be heard. For example, our findings indicate that the service may be particularly valuable for families with prior maltreatment histories, while most policy discussions have focused only on primary prevention.

Implications for research. The primary implication of this study for home visiting research is that it is possible to examine key research questions within the context of a flexible, community-based home visiting program that is unlike other researcher-developed program models. The findings related to program engagement and maltreatment outcomes are consistent with controlled studies indicating that many of the same challenges exist whether one is delivering services from the auspices of a university-based research trial with incentives to participate or from a grassroots agency struggling to keep families engaged based largely on their reputation in the community and each nurse-client iteration.

A key finding from this study was the difference in maltreatment report rates for families who enroll prenatally compared to those who begin in the postpartum period. This is a unique finding as the dominant nurse home visiting model does not offer the flexibility of beginning visits after the birth of the child. The findings support the notion that the prenatal period may be the best time to intervene, but also raise the question of how to develop effective services for those who do not connect with services during this period. Home visitation is a voluntary service in most cases. The answer may not be to make sure all families receive prenatal visits but that the timing of these visits and the program delivered is best suited to the needs of the family. Further research is needed to systematically study the difference in outcomes for families who engage at different times to optimize outcomes.

The findings related to parity build on a rather weak body of literature to better understand the needs and outcomes for first-time mothers versus those who already have children. This study clearly indicates that mothers who already have children are at least the same level of risk, if not higher, than first time mothers. The level of engagement with services was identical despite the number of children; this suggests that the demand exists for mothers who have children. Given that the bulk of the nurse home visiting findings are from a model that only serves first-time mothers, more research is needed to examine the best way to provide services in the context of a family with other children.

Another sparse area of research that this study contributes to is the provision of services in rural communities. As states expand services into previously unserved areas through MIECHV, more rural areas will be reached by home visiting services than before. However, there is little research that focuses on differences in the effectiveness of home visiting services in a rural setting or the predictors of successful implementation of services in these areas (Sweet &

Applebaum, 2004). Interventions are often developed in urban settings and tested on available populations often with the support of an existing service infrastructure. Thus, rural communities are typically the last to be tested for generalizability.

Rural areas can present challenges as well as strengths in supporting home visiting strategies. Challenges include poverty, barriers to client engagement due to issues of lack of trust of an outside system, and lack of access to needed support services, particularly mental health services (Silovsky, Bard, Chaffin, Hecht, Burris, Owora, Beasley, Doughty, & Lutzker, 2011). Rural mothers are also more likely report delayed or no prenatal care partly due to the lack of available local care increasing the risk for preterm delivery and infant mortality (Gamm, et al., 2003; Abma, et al., 1997). However, the close-knit community support including faith-based organizations and the local expertise of established community service providers can be strengths (Silovsky, et al., Lambert, Donahue, Mitchell, & Strauss, 2001). However, few studies have examined outcomes among populations in rural communities.

There is a need to increase research on different approaches to home visitation. As programs "scale-up" across states, an intensive area of focus for home visiting research is on the dissemination and implementation of evidence-based models. There is much concern about intervention fidelity, training and supervision of visitors, culture, climate, and other organizational factors that may enhance or impede the effectiveness of a previously tested intervention. This program does not have a standardized intervention model in the sense that visits are highly structured. The nurse is given flexibility to serve the family around a set of comprehensive clinical guidelines and regular supervision. While this is viewed as a potentially negative factor for an "evidence-based" program due to issues of replicability, this may be of intrinsic value to the success of the model. More research is needed in programs such as NFN to

determine if a balance between structure and clinical expertise can be adequately measured and replicated with fidelity.

There were several challenges that this study faced with regards to methodology that can inform future research. Measurement of treatment dosage is a continual struggle in this type of research. Utilizing multiple measures of service utilization (number of weeks, visits, hours, skills delivered, etc.) provide a more complete picture of the complicated delivery of home visiting services. It is also important to measure both initial engagement as well as ongoing retention in the program as these constructs appear to have different predictors. As a long-term intervention, ongoing participation in services is critical. This study also contributed a new measure of risk created uniquely for this study using existing agency records. The use of a well-constructed cumulative risk score that addresses multiple domains of both maternal and infant risk can have high predictive utility.

Implications for policy. The current policy landscape is one of expansion and evaluation of home visiting services. States vary as to the stage of development of their early childhood service system, but in most cases, an early childhood home visiting program is at the center of this system. States are currently struggling with deciding where home visiting will be expanded, for whom, and which model(s) will be provided. This is an opportunity for experimentation in services. There is currently an incentive, both financial and political, to only utilize those models that have demonstrated effectiveness with randomized trials. However, with a closer look at the literature, it is quite clear that there are no program models that have been proven to be effective for all families. Particularly in the case of maltreatment prevention, there is certainly no program model that is truly "effective". The opportunity to invest in "promising

approaches" such as NFN, which have been developed for the community which they serve, might be a viable option.

The most "mature" home visiting program in the United States is Nurse-Family Partnership. NFP's popularity and support has been bolstered by the rigorous and extensive longitudinal evaluation of outcomes (Scribano, 2010; MacMillan, 2009). The goals of NFP are quite ambitious, seeking to alter the "adverse maternal health-related behaviors during pregnancy, compromised care of the child, and stressful conditions in families' homes" and thereby prevent "the most pervasive and intractable problems faced by young children and parents" (Olds, et al., 1999). On the other hand, there are many populations that are not served by this model and some of the components needed for effective service remain open to debate.

Results of NFP analyses have consistently concluded that the following four factors are essential to the success of home visiting: home visitors must be nurses (not paraprofessionals), services should be targeted to at-risk families (not delivered universally), program protocols should consistent of evidence-based clinical methods, and fidelity to the tested model must be adhered in practice (Olds, et al, 1999). The rationale for selecting nurses was due to their healthcare training and also "their competence in managing the complex clinical situations often presented by at-risk families" and "increased credibility and persuasive power in the eyes of family members" (Olds, et al., 1999, p. 49). The results of the Denver trial found that paraprofessional benefits overall were about half the magnitude of those produced by nurses (Gomby, 2005). It is not clear, however, that other professionals might be equally effective.

Olds (2008) has suggested that the results of NFP clearly indicate that the benefits of home visitation are greatest for low-income, unmarried women who are at greater risk for welfare dependence, substance abuse, and crime. The corollary to this point is that given finite

resources, universal delivery of this type of intensive program to lower risk families would be "wasteful" and lead to a "dilution of services" (Olds, 2008; Olds, et al., 1999). This conclusion stands in stark contrast to the recommendation of the US Advisory Board on Child Abuse and Neglect (Krugman, 1993) that specifically decided to recommend a universal approach due to the desire to avoid stigmatizing the program and the broad appeal of the health, educational, and welfare benefits of the program to all families. While the concerns of program cost are certainly valid, expanding the reach and acceptability of this type of voluntary program could have immeasurable long-term outcomes to society as whole by shifting the culture of parenting towards openness to outside help.

Implication for practice. While study findings may be beneficial for home visiting practice at-large, this section focuses on how findings can be applied to the NFN agency and improve the quality of their current practice. Part of the benefit in engaging in evaluative research is the opportunity to examine what appears to be working and what areas need rethinking. This is something both practitioners and program administrators must consider. In many ways, this is the sign of an "evidence-based" program, one that is willing to honestly and openly use research and empiricism to improve their own practice and share their findings with others to improve practice more broadly.

One area of focus should be program engagement. The primary reasons for studying parental engagement and participation is to better understand the complexity of home visiting programs and to guide service improvement (Korfmacher, et al., 2008). If the agency hopes to impact the types of long-term outcomes that it is capable and positioned to address, there must be a renewed interest in improving the relationship between the client and the visitor so that initial engagement and retention rates improve. This improvement may begin with examining the

current culture of the agency. Perhaps nurses do not necessarily see themselves as a semi-permanent fixture in the lives of the families but as a crisis intervener that should provide a minimal level of services needed to stabilize the situation. This may lead to a different approach to developing a relationship with the caregiver.

A review of engagement strategies in parent and child mental health services (Ingoldsby, 2010) found programs utilize a variety of techniques to improve engagement. However, the review found only seventeen randomized trials in thirty years of literature. Based on these studies, approaches that demonstrated the most success were brief early engagement discussions, family systems approaches, enhancing family support, and motivational interviewing. Across all of these approaches, the provider directly addressed the issue of engagement with the family by identifying program benefits, discussing family expectations for outcomes, and working with the family to address potential barrier, both practical and psychological (Ingoldsby, 2010).

There is also a great deal of evidence that brief behavioral interventions such as motivational interviewing (Miller, 1983) can be quite effective in improving engagement across a diverse set of public health and social service programs and interventions (Dunn, Deroo, & Rivara, 2001; Lundahl, et al., 2010). Motivational interviewing was originally developed to help clients engage in substance abuse services by first addressing barriers to behavior change and reducing ambivalence around the target behavior. These same strategies have been successfully applied to improve retention in parent-training interventions including Parent-Child Interaction Therapy (Chaffin, et al., 2004, 2009), parent management training (Nock & Kazdin, 2005), and SafeCare (Damashek, et al., 2011).

On the other hand, cost and demand for services as well as respect for client time suggest that a program should be designed to tailor the dose and type of services to a family's unique set

of needs, but no more. Indeed, this idea of providing a "minimally sufficient" level of service is a central component to the Triple P program's universal parenting program (Sanders, 1999). While essentially all families served by NFN could be considered "high-risk", prescribing a level and type of service at the outset of services might be one way to improve initial engagement from families and commitment to complete the agreed-upon service plan. Of course, this would require careful mapping of initial risk to program components.

The second major implication of this research for practice is around the timing of program enrollment in the mother's life. Perhaps the more positive outcomes for prenatally enrolled families are related to an unmeasured characteristic of the families, or it could be an indication that the visitors do a better job of engaging and working with pregnant women. One option would be to become more aggressive about prenatal recruitment, but a second option may be to look more closely at the postnatal group, the client-nurse relationships and the model fit with the needs of these families to bring the outcomes more in line with those enrolled prenatally.

Another key area for improving practice is a need to focus more on addressing the mental health needs of the clients served by NFN. Although mental health issues are often addressed secondarily in home visiting services, recent literature suggests that given the high prevalence of mental health problems of mothers in these programs (29-60%; Ammerman, Putnam, Bosse, Teeters, & Van Ginkel, 2010), home visiting services may provide a promising venue to identify and treat maternal mental health symptoms using evidence-based approaches (Ammerman, Putnam, Chard, Stevens, Van Ginkel, 2011). Currently, the agency is exploring ways to integrate this into nurse home visitation.

Moving forward, the expansion of evidence-based home visiting program models must be accompanied by field studies that re-imagine the possibility of home visitation with today's modern family. The "technology" of home visiting is really decades, if not centuries, behind the times. Someone observing a home visit today and one in the 1980s or the 1940s may have a difficult time noticing much difference in the core content of the visits. This means incorporating and testing more technology including web-based social networking interfaces. As home visiting is widely disseminated to new families, we must use apply knowledge from health communications research to tailor the messages of home visiting to directly reach each individual family (Nansel, Weaver, Jacobsen, Glasheen, & Kreuter, 2008). There is also more thinking about how to use information technology and data management systems for quality improvement while integrating multiple data systems (Falconer, Rhodes, Mena, & Reid, 2009; McCabe, Potash, Omohundro, & Taylor, 2012). The use of information and internet technology should not replace the key "technology" of home visiting (the visitor-caregiver relationship), but should be used to augment it. The internet and new media are capable of bringing people together in new ways, for longer periods of time, and across further geographic distances than were possible before. Focus groups recently responded favorably to an online program for evidence-based parenting noting the "importance of a sense of community and learning through others' experiences" (Love, et al., 2013, p. 20).

Conclusion

This study confirmed two findings consistently produced in studies of early childhood home visiting. First, the caregivers served by this program are attempting to make the transition to life with a new baby while also dealing with an array of bitterly challenging life situations. The nurses who work alongside these families provide support and solutions for a very complex

set of obstacles. Forming and sustaining an engaged, participatory relationship between the agency and the mother is likely the hardest, but also the most valuable, part of this work. For home visiting to be effective, home visits must occur over a sustained period of time. Second, the evidence presented for whether or not NFN can be an effective means of preventing child abuse or neglect is mixed. This is consistent in many ways with the larger body of home visiting research. Some subgroups of families who receive services do appear to receive benefit. However, in most cases the likelihood that a family will come to the attention of the child welfare system is not improved.

The research questions for this study were developed to replicate questions asked of prior RCTs but using data from an existing community agency. This contribution provides additional support for the generalizability of findings that were consistent across study design types. This is important because the delivery context of NFN is likely to be similar to what the majority of families utilizing home visiting will receive, particularly as evidence-based models are widely disseminated under MIECHV expansion. Minimizing the research-practice gap is now of critical importance (Proctor, et al., 2009). Even with highly efficacious program models, successfully delivering and studying interventions "in context" has proven to be extremely challenging (Mendel, et al., 2008). While there will be some exceptions, it is likely that home visiting programs will largely be delivered via existing service channels in embedded mental health or community health care organizations, and not driven by or supported by university researchers conducting field studies. Further, like NFN, expansion home visiting programs will serve high-risk families in the local community. So, the fact that these study findings are largely in concert with findings from randomized-controlled trials is a strong contribution to a literature often lacking in studies that value generalizability, replication, and comparative effectiveness.

The downside to this coherence with prior literature is the failure to identify an effect on the prevention of child abuse and neglect, a "bottom-line benefit" of home visiting (Chaffin, 2004). While families in the prenatal group did appear to have a very low rate of later maltreatment, there is no way to be sure whether or not this is simply a selection effect. This study does provide some reasonable evidence that families who receive multiple postpartum visits receive no protective benefit from the program compared to those who dropout after one visit. Studies from Healthy Families trials have examined possible explanations for the lack of effectiveness in preventing maltreatment (Duggan, et al., 2004). Summarized by Chaffin (2004), the home visitors in these studies were not equipped to identify and manage the strongest risk factors for maltreatment (partner violence, substance abuse, and parental depression). The nurses delivering visits in the current study are specifically trained to at least screen for these concerns and make appropriate referrals. Although nurses may be equipped to identify and intervene, if families are not engaged for an adequate period of time, the window of opportunity may be not open long enough to realize change.

There is some evidence to the contrary, that greater retention is not the key to improved results (Landsverk, et al., 2002; Duggan, et al., 2004), suggesting that "problems with effectiveness may lie more with the model itself (Chaffin, 2004, p. 593). One home visiting model, Every Child Succeeds, is currently developing and implementing "augmented modules" to directly address known risk factors for maltreatment that fill a need in existing home visiting service models (Ammerman, et al., 2007). For example, their Maternal Depression Module including In-Home Cognitive Behavioral Therapy has shown initial evidence of success (Ammerman, et al., 2007). It may be these types of "evidence-based" changes or adaptations to

existing home visiting models that leads to making a real difference in maltreatment prevention outcomes.

Home visiting services represent an investment in the life of a child made by the caregiver and society. Evidence is mounting that intervening in the early childhood period is best because individuals are most vulnerable and there is also the most potential for positive gains to be made. Support is growing for this notion and social work research, policy, and practice must be prepared to meet this demand for action.

References

Abma J., Chandra A., Mosher W., Peterson L., & Piccinino L. (1997). Fertility, family planning, and women's health: New data from the 1995 National Survey of Family Growth. National Center for Health Statistics. *Vital Health Statistics*. Series 23, Number 19.

Abrams, L. S. & Curran, L. (2009). "And you're telling me not to stress?" A grounded theory study of postpartum depression symptoms among low-income mothers. *Psychology of Women Quarterly, 33*(3), 351–362.

Ackerman, P. T., Newton, J. E., McPherson, W. B., Jones, J. G., & Dykman, R. (1998). Prevalence of post-traumatic stress disorder and other psychiatric diagnoses in three groups of abused children (sexual, physical, and both). *Child Abuse & Neglect, 22*(8), 759-74.

Addy, S., & Wight, V. (2012). *Basic facts about low-income children, 2010. Children under age 3.* New York, NY: National Center for Children in Poverty, Columbia University, Mailman School of Public Health.

Administration for Children, Families. (2002). Making a difference in the lives of children and families: The impacts of Early Head Start Programs on young children and their families. Summary report. Washington, DC: DHHS.

Ainsworth, M. D., Blehar, M., Waters, E., & Wall, S. (1978). *Patterns of Attachment: A Psychological Study of the Strange Situation*. Hillsdale: Erlbaum.

Ainsworth, M., & Bowlby, J. (1991). An ethological approach to personality development. *American Psychologist*, *46*, 331-341.

Allison, P. (1995). Survival analysis using the SAS System: A practical guide. Cary, NC: SAS Institute.

American Academy of Pediatrics. (2001). Developmental surveillance and screening of infants and young children. *Pediatrics, 108,* 92-95.

Ammerman, R. T., Stevens, J., Putnam, F. W., Altaye, M., Hulsmann, J. E., Lehmkuhl, H. D., Monroe, J. C., Gannon, T.A., & Van Ginkel, J. B. (2006). Predictors of Early Engagement in Home Visitation. *Journal of Family Violence*, *21*(2), 105–115.

Ammerman, R., Putnam, F., Chard, K., Stevens, J., & Van Ginkel, J. (2011). PTSD in depressed mothers in home visitation. *Psychology Trauma: Theory, Research, Practice, & Policy*. Advance online publication.

Anderson, L., Shinn, C., Fullilove, M., Scrimshaw, S., Fielding, J., & Normand, J. et al. (2003). The effectiveness of early childhood development programs. A systematic review. *American Journal of Preventive Medicine, 24*, 32-46.

Aos, S., Lieb, R., Mayfield, J., Miller, M., & Pennucci, A. (2004). *Benefits and costs of prevention and early intervention programs for youth*. Olympia, WA: Institute for Public Policy.

Appleyard, K., Egeland, B., Dulmen, M. H., & Alan Sroufe, L. (2005). When more is not better: The role of cumulative risk in child behavior outcomes. *Journal of Child Psychology and Psychiatry*, *46*(3), 235-245.

Appleyard, K., Berlin, L., Rosanbalm, K., & Dodge, K. (2011). Preventing early child maltreatment: Implications from a longitudinal study of maternal abuse history, substance use problems, and offspring victimization. *Prevention Science, 12*, 139-149.

Ashdown-Lambert, J. R. (2005). A review of low birth weight: predictors, precursors and morbidity outcomes. *The Journal of the Royal Society for the Promotion of Health*, *125*(2), 76-83.

Astuto, J. & Allen, L. (2009). Home visitation and young children: An approach worth investing in? *Social Policy Report, 23*(4), 3-21.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46*(1), 399-424.

Bae, H., Solomon, P. L. & Gelles, J. (2008). Multiple child maltreatment recurrence relative to single recurrence and no recurrence. *Children and Youth Services Review, 31*, 617-624.

Bagnato, S., Suen, H., Brickley, D., Smith-Jones, J., Dettore, E. (2002). Child developmental impact of Pittsburgh's Early Childhood Initiative (ECI) in high-risk communities: First-phase authentic evaluation research. *Early Childhood Research Quarterly*, *17*, 559-580.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.

Bandura, A. (1994). Self-Efficacy. In V. Ramachaudran (Ed.), *Encyclopedia of human behavior* (pp. 71-81). New York: Academic Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W.H. Freeman.

Barlow, J., Simkiss, D., & Stewart-Brown, S. (2006). Interventions to prevent or ameliorate child physical abuse and neglect: findings from a systematic review of reviews. *Journal of Children's Services, 1*(3), 6-28.

Barnett, W. (2000) Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of Children, 5*, 25-50.

Barth, R. P., Hacking, S., & Ash, J. (1988). Preventing child abuse: An experimental evaluation of the child parent enrichment project. *The Journal of Primary Prevention, 8*(4), 201-217.

Belsky, J. (1980). Child maltreatment: An ecological integration. *American Psychologist*, *35*(4), 320-335.

Belsky, J. (1984). The determinants of parenting: A process model. *Child Development*, 83-96.

Belsky, J. (1993). Etiology of child maltreatment: A developmental-ecological analysis. *Psychological Bulletin*, *114*(3), 413.

Berkoff, M. C., Leslie, L. K., & Stahmer, A. (2006). Accuracy of caregiver identification of developmental delays among young children involved with child welfare. *Journal of Developmental and Behavioral Pediatrics*, *27*(4), 310-318.

Berlin, L., Appleyard, K., & Dodge, K. (2011). Intergenerational continuity in child maltreatment: Mediating mechanisms and implications for prevention. *Child Development*, *82*(1), 162-176.

Berrick, J., Needell, B., Barth, R., & Jonson-Reid, M. (1998).*The tender years: Toward developmentally-sensitive child welfare services for very young children.* New York: Oxford University Press.

Bilukha, O., Hahn, R., Crosby, A., Fullilove, M. T., Liberman, A., Moscicki, E., & Snyder, S., (2005). The effectiveness of early childhood home visitation in preventing violence: a systematic review. *American Journal of Preventive Medicine*, *28*(2 Suppl 1), 11-39.

Boonstra, H. D. (2009). Home Visiting for At-Risk Families: A Primer On a Major Obama Administration Initiative. *Guttmacher Policy Review*, *12*(3), 11-15.

Boris, N., Larrieu, J., Zeanah, P., Nagle, G., Steier, A., & McNeill, P. (2006). The process and promise of mental health augmentation of nurse home-visiting programs: Data from the Louisiana Nurse-Family Partnership. *Infant Mental Health Journal*, *27*(1), 26-40.

Boulatoff, C., & Jump, V. K. (2007). Blueprint of a cost analysis approach for early intervention: Application to a home visiting program to prevent child abuse and neglect. *Journal of Early Intervention*, *30*(1), 73-84.

Bowlby, J. (1969). *Attachment and loss, Vol. 1: Attachment.* New York: Basic Books.

Bowlby, J. (1973). *Attachment and loss, Vol. 2: Separation.* New York: Basic Books.

Bowlby, J. (1980). *Attachment and loss, Vol. 3: Loss, sadness and depression.* New York: Basic Books.

Boyd, R. C., Le, H. N., & Somberg, R. (2005) Review of screening instruments for postpartum depression. *Archives of Women's Mental Health, 8,* 141–53

Brayden, R. (1993). A prospective study of secondary prevention of child maltreatment. *The Journal of Pediatrics*, *122*(4), 511-516.

Bretherton, I. (1992). The origins of attachment theory: John Bowlby and Mary Ainsworth. *Developmental Psychology*, *28*(5), 759-775.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*(7), 513-531.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge: Harvard University Press.

Brooks-Gunn, J., & Duncan, G. J.(1997). The Effects of Poverty on Children. *Future of Children*, *7*(2), 55–71.

Brown, C. H. & Sturgeon, S. (2004). Promoting a healthy start in life and reducing early risks. In C. Hosman, E. Jané-Llopis & S. Saxena (Eds). *Prevention of mental disorders: effective interventions and policy options: Summary report.* World Health Organization Dept. of Mental Health and Substance Abuse in collaboration with the Prevention Research Centre of the Universities of Nijmegen and Maastricht.

Browne, K., & Parr, R. (1980). Contributions of an ethological approach to the understanding of abuse. In N. Frude (Ed.), *Psychological approaches to child abuse* (pp. 83-99). London: Batsford Academic and Educational.

Bruskas, D. (2008). Children in foster care: A vulnerable population at risk. *Journal of Child and Adolescent Psychiatric Nursing*, *21*(2), 70-77.

Burwick, A., Strong, D., Xue, Y., Koball, H., & Coffee-Borden, B. (2011). *Supporting Evidence-based home visiting to prevent child maltreatment: cross-site evaluation cost study background and design update*. Washington, DC: Children's Bureau, Administration for Children and Families, US DHHS.

Caldera, D., Burrell, L., Rodriguez, K., Crowne, S. S., Rohde, C., & Duggan, A. (2007). Impact of a statewide home visiting program on parenting and on child health and development. *Child Abuse & Neglect*, *31*(8), 829-852.

Campbell, F. & Ramey, C. (1994). Effects of early intervention of intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development, 65*, 684-698.

Casanueva, C., Cross, T., Ringeisen, H., & Christ, S. (2011). Prevalence, trajectories, and risk factors for depression among caregivers of young children involved in child maltreatment investigations. *Journal of Emotional and Behavioral Disorders, 19*(2), 98-116.

Chaffin, M., Silovsky, J. F., Funderburk, B., Valle, L. A., Brestan, E. V., Balachova, T., & Bonner, B. L. (2004). Parent-child interaction therapy with physically abusive parents: Efficacy for reducing future abuse reports. *Journal of Consulting and Clinical Psychology*, *72*(3), 500-510.

Chaffin, M., Valle, L. A., Funderburk, B., Gurwitch, R., Silovsky, J., Bard, D., & Kees, M. (2009). A motivational intervention can improve retention in PCIT for low-motivation child welfare clients. *Child Maltreatment, 14*(4), 356-368.

Chaffin, M., & Bard, D. (2006). Impact of intervention surveillance bias on analyses of child welfare report outcomes. *Child Maltreatment, 11*(4), 301-312.

Chaffin, M., Kelleher, K., & Hollenberg, J. (1996). Onset of physical abuse and neglect: Psychiatric, substance abuse, and social risk factors from prospective community data. *Child Abuse & Neglect, 20*(3), 191-203.

Chaudron, L. H., Kitzman, H. J., Peifer, K. L., Morrow, S., Perez, L. M., & Newman, M. C. (2005). Self-recognition of and provider response to maternal depressive symptoms in low-income Hispanic women. *Journal of Women's Health, 14*, 331-338.

Chen, E., Matthews, K., & Boyce, W. T. (2002). Socioeconomic differences in children's health: How and why do these relationships change with age? *Psychological Bulletin*, *128*(2), 295–329.

Child Welfare Information Gateway. (2008). *Preventing child abuse and neglect factsheet.* Washington, D.C.: Children's Bureau, Administration for Children, Youth, and Families.

Cicchetti, D. & Rogosch, F. (2001). Diverse patterns of neuroendocrine activity in maltreated children. *Development and Psychopathology, 13*, 677-693.

Cicchetti, D., & Carlson, V. (Eds.). (1989). *Child maltreatment: Theory and research on the causes and consequences of child abuse and neglect*. New York: Cambridge University Press.

Cicchetti, D., & Lynch, M. (1993). Toward an ecological/transactional model of community violence and child maltreatment: Consequences for children's development. *Psychiatry*, *56*, 96-188.

Cicchetti, D., & Rizley, R. (1981). Developmental perspectives on the etiology, intergenerational transmission and sequelae of child maltreatment. *New Directions for Child Development* (pp. 31-56). San Francisco: Jossey-Bass.

Cicchetti, D., & Toth, S. (1995). A developmental psychopathology perspective on child abuse and neglect. *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*(5), 541-565.

Cicchetti, D., & Toth, S. L. (2005). Child maltreatment. *Annual Review of Clinical Psychology*, *1*, 409-38.

Coca-Perraillon, M. (2007). Local and global optimal propensity score matching. *SAS Global Forum 2007, Statistics and Data Analysis.* Paper 185-2007.

Colman, R., & Widom, C. S. (2004). Childhood abuse and neglect and adult intimate relationships: a prospective study. *Child Abuse & Neglect*, *28*(11), 1133-51.

Conron, K. J., Beardslee, W., Koenen, K. C., Buka, S. L., & Gortmaker, S. L. (2009). A longitudinal study of maternal depression and child maltreatment in a national sample of families investigated by child protective services. *Archives of Pediatrics & Adolescent Medicine*, *163*(10), 922.

Constantino, J.N. (2000). The "primacy" of early experience. *Journal of Psychiatric Practice*, *6*, 235-236.

Corso, P. S., & Fertig, A. R. (2010). The economic impact of child maltreatment in the United States: are the estimates credible? *Child Abuse & Neglect*, *34*(5), 296-304.

Corso, P. S., & Lutzker, J. R. (2006). The need for economic analysis in research on child maltreatment. *Child Abuse & Neglect*, *30*(7), 727-38.

Corso, P. S., Edwards, V. J., Fang, X., & Mercy, J. a. (2008). Health-related quality of life among adults who experienced maltreatment during childhood. *American Journal of Public Health*, *98*(6), 1094-100.

Corso, P. & Filene, J. H. (2009). Programmatic cost analysis of the Family Connections program. *Protecting Children*, *24*(3), 78-88.

Coulton, C. J., Korbin, J. E., & Su, M. (1999). Neighborhoods and child maltreatment: a multi-level study. *Child Abuse & Neglect*, *23*(11), 1019-40.

Coulton, C. J., Crampton, D. S., Irwin, M., Spilsbury, J. C., & Korbin, J. E. (2007). How neighborhoods influence child maltreatment: a review of the literature and alternative pathways. *Child Abuse & Neglect*, *31*(11-12), 1117-42.

Courtney, M. (1999). The economics. *Child Abuse & Neglect*, *23*(10), 975-986.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society*, *34*(2), 187-220.

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *The British Journal of Psychiatry*, *150*(6), 782-786.

Currie, J., & Spatz Widom, C. (2010). Long-term consequences of child abuse and neglect on adult economic well-being. *Child Maltreatment*, *15*(2), 111-120.

D'Agostino, R. (1998). Tutorial in biostatistics: Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine*, *17*, 2265-2281.

Damashek, A., Doughty, D., Ware, L., & Silovsky, J. (2011). Predictors of client engagement and attrition in home-based child maltreatment prevention services. *Child Maltreatment, 16*(1), 9-20.

Daro, D. & Dodge, K. (2010). Strengthening home-visiting intervention policy: Expanding research, building knowledge. In Haskins, R. & Barnett (Eds). *New Directions for America's Preschool Policies*. Washington D.C.: NIERR and Brookings (pp 79-86).

Daro, D. (1993). Child maltreatment research: Implications for program design. In D. Cicchetti & S. L. Toth (Eds.), *Child abuse, child development, & social policy* (pp. 331-367). New York: Ablex Publishing.

Daro, D. (2009). *Embedding home visitation programs within a system of early childhood services*. Chicago. Chapin Hall at the University of Chicago.

Daro, D. (2012). Crafting effective child abuse prevention systems: A legacy of vision. In Krugman, R. & Korbin, J. (Eds) (forthcoming). *C. Henry Kempe: A 50 year legacy to the field of child abuse and neglect*. New York: Springer Publishing.

Daro, D., Dodge, K., Weiss, H., & Zigler, E. (2009). Home visiting initiative. A letter to President Obama.

Daro, D., McCurdy, K., Falconnier, L., & Stojanovic, D. (2003). Sustaining new parents in home visitation services: key participant and program factors. *Child Abuse & Neglect, 27*, 1101-1125.

Dawes Knight, E., Smith, J. B., Dubowitz, H., Litrownik, A. J., Kotch, J. B., English, D., & Everson, M. D., (2006). Reporting participants in research studies to Child Protective Services: limited risk to attrition. *Child Maltreatment, 11*(3), 257-62.

Dawson, P., Van Doorninck, W., & Robinson, J. (1989). Effects of home-based, informal social support on child health. *Journal of Developmental and Behavioral Pediatrics, 10*, 63-67.

De La Rosa, I., Perry J., & Johnson V. (2009) Benefits of increased home-visitation services: exploring a case management model. *Family and Community Health, 32*(1), 58-75.

DeBellis, M. D. (2001). Developmental traumatology: the psychobiological development of maltreated children and its implications for research, treatment, and policy. *Development and Psychopathology*, *13*(3), 539-64.

DeBellis, M. D. (2007). *Psychobiology of neglect*. Presentation at Translational Research on Child Neglect Consortium, Raleigh, NC: Duke University .

DePanfilis D., Zuravin S.J. (1998). Rates, patterns and frequency of child maltreatment recurrences among families known to CPS. *Child Maltreatment, 3*, 27–42.

DePanfilis, D., & Zuravin, S. J. (2002). The effect of services on the recurrence of child maltreatment. *Child Abuse & Neglect*, *26*(2), 187-205.

Diez-Roux, A. V. (2000). Multilevel analysis in public health research. *Annual Review of Public Health*, *21*, 171-192.

Doll, L. S., Saul, J. R., & Elder, R. W. (2007). Injury and violence prevention interventions : An Overview. In L. Dolls, S. Bonzo, D. Sleet, & E. Mercy (Eds.), *Handbook of injury and violence prevention* (pp. 21-33).

Drake, B., Jonson-Reid, M., Way, I., & Chung, S. (2003). Substantiation and recidivism. *Child Maltreatment*, *8*(4), 248-260.

Drake, B., Jolley, J. M., Lanier, P., Fluke, J., Barth, R. P., & Jonson-Reid, M. (2011). Racial bias in child protection? A comparison of competing explanations using national data. *Pediatrics*, *127*, 471-478.

Drake, B., Jonson-Reid, M., & Sapokaite, L. (2006). Re-reporting of child maltreatment: Does participation in other public sector services moderate the likelihood of a second, maltreatment report? *Child Abuse & Neglect, 30*(11), 1201.

Drake, B., Jonson-Reid, M., Way, I., & Chung, S. (2002). Recidivism in Child Protective Services among Substantiated and Unsubstantiated Cases. *Washington: Office of Child Abuse and Neglect, Department of Health and Human Services*.

Duggan, A., Berlin, L., Cassidy, J., Burrell, L., & Tandon, D. (2009). Examining maternal depression and attachment insecurity as moderators of the impacts of home visiting for at-risk mothers and infants. *Journal of Consulting and Clinical Psychology, 77*(4), 788-799.

Duggan, A., Caldera, D., Rodriguez, K., Burrell, L., Rohde, C., & Crowne, S. S. (2007). Impact of a statewide home visiting program to prevent child abuse. *Child Abuse & Neglect, 31*(8), 801-27.

Duggan, A., McFarlane, E., Fuddy, L., Burrell, L., Higman, S. M., Windham, A., & Sia, C. (2004). Randomized trial of a statewide home visiting program: impact in preventing child abuse and neglect. *Child Abuse & Neglect, 28*(6), 597-622.

Duggan, A., Windham, A., McFarlane, E., Fuddy, L., Rohde, C., Buchninder, S., & Sia, C. (2000). Hawaii's Healthy Start Program of home visiting for at-risk families: Evaluation of family identification, family engagement, and service delivery. *Pediatrics, 105*, 250-259.

DuMont, K., Mitchell-Herzfeld, S., Greene, R., Lee, E., Lowenfels, A., Rodriguez, M., & Dorabawila, V. (2008). Healthy Families New York (HFNY) randomized trial: Effects on early child abuse and neglect. *Child Abuse & Neglect, 32*(3), 295-315.

DuMont, K., Kirkland, K., Mitchell-Herzfeld, S., & Ehrhard-Dietzel, S. (2010). *Final Report: A randomized trial of Healthy Families New York (HFNY): Does home visiting prevent child maltreatment?* New York State Office of Children and Family Services.

Duncan, G. J., Brooks-Gunn, J., & Klebanov, P. K. (1994). Economic deprivation and early childhood development. *Childhood Development, 65*(2), 296–318.

Dunn, C., Deroo, L., & Rivara, F. P. (2001). The use of brief interventions adapted from motivational interviewing across behavioral domains: a systematic review. *Addiction*, *96*(12), 1725-1742.

Easterbrooks, M. A., Jacobs, F. H., Bartlett, J. D., Goldberg, J., Contreras, M. M., Kotake, C., Raskin, M., & Chaudhuri, J. H. (2013). Initial findings from a randomized, controlled trial of Healthy Families Massachusetts: Early program impacts on young mothers' parenting. Report published by Pew Center on the States. Retrieved from http://www.pewstates.org/uploadedFiles/PCS_Assets/2013/Healthy_Families_Massachusetts_report.pdf

Eberhard-Gran, M., Eskild, A., Tambs, K., Opjordsmoen, S., & Samuelsen, S. O. (2001). Review of validation studies of the Edinburgh Postnatal Depression Scale. *Acta Psychiatrica Scandinavica*, *104*(4), 243–9.

Egeland, B., Jacobvitz, D., & Sroufe, A. (1988). Breaking the Cycle of Abuse. *Child Development, 59*, 1080-1088.

Egeland, B., Yates, T., Appleyard, K., & van Dulmen, M. (2002). The long-term consequences of maltreatment in the early years: A developmental pathway model to antisocial behavior. *Children's Services*, *5*(4), 249-260.

Emerson, E. & Hatton, C. (2007). Mental health of children and adolescents with intellectual disabilities in Britain. *The British Journal of Psychiatry, 191,* 493-499

Enders, C. K. (2010). *Applied missing data analysis*. The Guilford Press.

Engle, P. L., & Black, M. M. (2008). The effect of poverty on child development and educational outcomes. *Annals of the New York Academy of Sciences*, *1136*(1), 243-256.

Evans, G. W. (2003). A multimethodological analysis of cumulative risk and allostatic load among rural children. *Developmental Psychology, 39*(5), 924-933.

Falconer, M. K., Rhodes, T., Mena, K. C., & Reid, R. (2009). Management information systems: Applications in home visiting programs designed to prevent child abuse and neglect. *Journal of Technology in Human Services*,*27*(3), 194-215.

Fergusson, D. M., Grant, H., Horwood, L. J., & Ridder, E. M. (2005). Randomized trial of the Early Start program of home visitation. *Pediatrics, 116*(6), e803-9.

Filene, J. (2012). *Meta-analytic review of components associated with home visiting programs: Final report*. Report prepared for Pew Center on the States. James Bell Associates.

Finkelhor, D. (2009). The prevention of childhood sexual abuse. *The Future of Children, 19*(2), 169-94.

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behavior: An introduction to theory. Reading, MA: Addison-Wesley.

Fishbein, M., Guenther-Gray, C., Johnson, W., Wolitski, R., McAlister, A., Rietmeijer, C., et al. (1997). Using a theory- based community intervention to reduce AIDS risk behaviors: The CDC's AIDS community demonstration projects. In M. Goldberg, M. Fishbein, & S. Middlestadt (Eds.), Social marketing: Theory and practical perspectives (pp. 123–146). Mahwah, NJ7 Lawrence Erlbaum.

Folland, S., Goodman, A., & Stano, M. (2010). *The economics of health and health care* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Foster, E. M., Johnson-Shelton, D., & Taylor, T. K. (2007). Measuring time costs in interventions designed to reduce behavior problems among children and youth. *American Journal of Community Psychology*, *40*(1-2), 64-81.

Fraser, J., Armstrong, K. L., Morris, J. P., & Dadds, M. R. (2000). Home visiting intervention for vulnerable families with newborns: follow-up results of a randomized controlled trial. *Child Abuse & Neglect*, *24*(11), 1399-429.

Freed, R., Chan, P., Dingman, K., and Boger, M. (2012). Enhancing maternal depression recognition in health care settings: A review of strategies to improve detection, reduce barriers, and reach mothers in need. *Families, System, & Health, 30* (1), 1-18.

Fromm, S. (2001). *Total estimated cost of child abuse and neglect in the United States: Statistical evidence.* Prevent Child Abuse America.

Gaines, R., Sandgrund, A., Green, A. H., & Power, E. (1978). Etiological factors of maltreatment: A multivariate study of abusing, neglecting, and normal mothers. *Journal of Abnormal Psychology*, *87*(5), 531-540.

Galano, J. & Huntington, L. (2012). Comparison of primiparous and multiparous mothers: Healthy Families program participation, outcomes, challenges, and adaptations. Report prepared for Pew Center for the States. Retrieved from http://www.pewstates.org/uploadedFiles/PCS_Assets/2013/Second_Time_Mothers_and_Beyond_report.pdf

Gamm, L., Hutchison, L., Dorsey, A., & Dabney, B. (2003). (Eds.). *Rural Healthy People* 2020: *A Companion Document to Healthy People* 2020. A report by the Southwest Rural

Health Research Center, School of Rural Public Health, Texas A&M University System Health Science Center to the Health Resources and Services. Administration's Office of Rural Health Policy, US Department of Health and Human Services.

Garbarino, J. (1977). The human ecology of child maltreatment: A conceptual model for research. *Journal of Marriage and Family*, *39*(4), 721-735.

Garner, A., Shonkoff, J.P., Siegel, B.S., et al. (2012) American Academy of Pediatrics Policy Statement. Early childhood adversity, toxic stress, and the role of the pediatrician: Translating developmental science into lifelong health. *Pediatrics*, *129*(1), e224-e231.

Geeraert, L., Van den Noortgate, W., Grietens, H., & Onghena, P. (2004). The effects of early prevention programs for families with young children at risk for physical child abuse and neglect: a meta-analysis. *Child Maltreatment*, *9*(3), 277-91.

Gil, D. (1970). *Violence against children: Physical child abuse in the United States*. Cambridge: Harvard University Press.

Gomby, D. S. (2005). *Home Visitation in 2005: Outcomes for Children and Parents.* Investment in Kids Working Paper 7, Committee for Economic Development, Invest in Kids Working Group.

Gordon, R. S. (1983). An operational classification of disease prevention. *Public Health Reports*, *98*(2), 107-109.

Groenwold, R. H., Donders, A. R. T., Roes, K. C., Harrell, F. E., & Moons, K. G. (2012). Dealing with missing outcome data in randomized trials and observational studies. *American Journal of Epidemiology*, *175*(3), 210-217.

Guo, S, Barth, R., & Gibbons, C. (2006). Propensity score matching strategies for evaluating substance abuse services for child welfare clients. *Children and Youth Services Review, 28*(4), 357-383.

Guo, S., & Fraser, M. W. (2010). *Propensity score analysis: Statistical methods and applications*. Los Angeles: Sage Publications.

Hackman, D., Farah, M. J., & Meaney, M. J. (2010). Socioeconomic status and the brain: mechanistic insights from human and animal research. *Nature Reviews. Neuroscience*, *11*(9), 651–9.

Hall, L., Gurley, D., Sachs, B., & Kryscio, R. (1991). Psychosocial predictors of maternal depressive symptoms, parenting attitudes, and child behavior in single-parent families. *Nursing Research, 40*, 214-220

Hall, L., Williams, C., & Greenberg, R. (1985). Supports, stressors, and depressive symptoms in mothers of young children. *American Journal of Public Health, 75*, 518-521.

Hammen, C., Shih, J., & Brennan, P. (2004). Intergenerational transmission of depression: Test of an interpersonal stress model in a community sample. *Journal of Consulting and Clinical Psychology, 72*(3), 511-522.

Hammond, W. R. (2003). Public Health and Child Maltreatment Prevention: The Role of the Centers for Disease Control and Prevention. *Child Maltreatment, 8*(2), 81-83.

Harlow, H., Harlow, M., & Hanson, E. (1963). The maternal affectional system of Rhesus monkeys. In H. Rheingold (Ed.), *Maternal Behavior in Mammals*. New York: Wiley.

Harwood, M. D. & Eyberg, S. M. (2004). Therapist verbal behavior early in treatment: Relation to successful completion of Parent-Child Interaction Therapy. *Journal of Clinical Child and Adolescent Psychology. 33*(3), 601-612.

Hatcher, J., Rayens, M. K., Peden, A. R., & Hall, L. A. (2008). Predictors of depression in low income African American single mothers. *Journal of Health Disparities Research and Practice, 2*(3), 89-108.

Haviland, A., Nagin, D. S., & Rosenbaum, P. R. (2007). Combining propensity score matching and group-based trajectory analysis in an observational study. *Psychological Methods, 12*(3), 247-67.

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling [White paper]. Retrieved from http://www.afhayes.com/public/process2012.pdf

Heckman, J. J. (2008). Role of income and family influence on child outcomes. *Annals of the New York Academy of Sciences, 1136*, 307-23.

Hill, J. L., Brooks-Gunn, J., & Waldfogel, J. (2003). Sustained effects of participation in an early intervention for low-birth-weight premature infants. *Developmental Psychology, 39*, 730–744.

Howard, K. S., & Brooks-Gunn, J. (2009). The role of home-visiting programs in preventing child abuse and neglect. *The Future of Children, 19*(2), 119-46.

Hurlburt, M. S., Barth, R. P., Leslie, L., Landsverk, J., & McCrae, J. (2007). Building on strengths: Current status and opportunities for improvement of parent training for families in child welfare. In R. Haskins, F. Wulczyn, & M.B. Webb (Eds), *Child protection: Using research to improve policy and practice* (pp. 81-106). Washington DC: Brookings Institute Press.

Hussey, J. M., Marshall, J. M., English, D. J., Knight, E. D., Lau, A. S., Dubowitz, H., & Kotch, J. B. (2005). Defining maltreatment according to substantiation: distinction without a difference? *Child Abuse & Neglect*, *29*(5), 479-92.

Ingoldsby, E. M. (2010). Review of interventions to improve family engagement and retention in parent and child mental health programs. *Journal of Child and Family Studies*, *19*(5), 629-645.

IOM (Institute of Medicine) and National Research Council (NRC). 2012. *Child maltreatment research, policy, and practice for the next decade: Workshop summary*. Washington, DC: The National Academies Press.

Jee, S. H., Szilagyi, M., Ovenshire, C., Norton, A., Conn, A.-M., Blumkin, A., & Szilagyi, P. G. (2010). Improved detection of developmental delays among young children in foster care. *Pediatrics*, *125*(2), 282-9.

Jensen, P. S., Hoagwood, K., & Trickett, E. J. (1999). Ivory tower or earthen trenches? Community collaborations to foster real-world research. *Applied Developmental Science*, *3*(4), 206-212.

Joffe, M. M.,& Rosenbaum, P. R. (1999). Invited commentary: Propensity scores. *American Journal of Epidemiology, 150*, 327–333.

Jonson-Reid, M., Drake, B., & Kohl, P. L. (2009). Is the overrepresentation of the poor in child welfare caseloads due to bias or need? *Children and Youth Services Review*, *31*(3), 422–427.

Jonson-Reid, M. (1998). Youth violence and exposure to violence in childhood. An ecological review. *Aggression and Violent Behavior*, *3*(2), 159-179.

Jonson-Reid, M. (2002). Exploring the relationship between child welfare intervention and juvenile corrections involvement. *The American Journal of Orthopsychiatry*, *72*(4), 559-76.

Jonson-Reid, M., Kohl, P. L., & Drake, B. (2012). Child and adult outcomes of chronic child maltreatment. *Pediatrics*, *129*(5), 839-845.

Jonson-Reid, M., Drake, B., Kim, J., Porterfield, S., & Han, L. (2004). A prospective analysis of the relationship between reported child maltreatment and special education eligibility among poor children. *Child Maltreatment*, *9*(4), 382-94.

Jotzo, M., & Poets, C.F. (2005). Helping parents cope with the trauma of premature birth: An evaluation of a trauma-preventive psychological intervention. Pediatrics, 115, 915–919.

Kadushin, A., McGloin, J., & Martin, J. (1981). *Child abuse: An interactional effect*. New York: Columbia University Press.

Karazsia, B. T., & van Dulmen, M. H. (2008). Regression models for count data: Illustrations using longitudinal predictors of childhood injury. *Journal of Pediatric Psychology*, *33*(10), 1076-1084.

Karoly, L., Kilburn, M., & Cannon, J. (2005). *Early childhood interventions proven results, future promise*. CA: RAND.

Kempe, C. H., Silverman, F. N., Steele, B. F., Droegemueller, W., & Silver, H. K. (1962). The battered-child syndrome. *The Journal of the American Medical Association*, *181*(1), 17-24.

Kendall-Tackett, K. (2002). The health effects of childhood abuse: four pathways by which abuse can influence health. *Child Abuse & Neglect*, *26*(6-7), 715-29.

Kendrick, D. (2000). Does home visiting improve parenting and the quality of the home environment? A systematic review and meta-analysis. *Archives of Disease in Childhood, 82*(6), 443-451.

Kersting, A., Dorsch, M., Wesselmann, U., Lüdorff, K., Witthaut, J., Ohrmann, P., & Arolt, V. (2004). Maternal posttraumatic stress response after the birth of a very low-birth-weight infant: Stress, fibromyalgia, and sleep. *Journal of Psychosomatic Research, 57*(5), 473-476.

Kim, J., Cicchetti, D., & Manly, J. T. (2009). Child maltreatment and trajectories of personality and behavioral functioning: Implications for the development of personality disorder. *Development and Psychopathology, 21*(3), 889-912.

King, P. (2012). Replicability of structural models of the Edinburgh Postnatal Depression Scale (EPDS) in a community sample of postpartum African American women with low socioeconomic status. *Archive of Women's Mental Health, 15*, 77-86.

Kitzman, H., Olds, D. L., Henderson, C. R., Hanks, C., Cole, R., Tatelbaum, R., McConnochie, K. M., et al. (1997). Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing. *Journal of the American Medical Association, 278*(8), 644-652.

Kitzman, J., Olds, D. L., Cole, R. E., Hanks, C., Anson, E., Arcoleo, K. J., Luckey, D. W., et al. (2010). Enduring effects of prenatal and infancy home visiting by nurses on children: follow-up of a randomized trial among children at age 12 years. *Archives of Pediatrics & Adolescent Medicine, 164*(5), 412-8.

Kohl, P. L., Jonson-Reid, M., & Drake, B. (2009). Time to Leave Substantiation Behind Findings From A National Probability Study. *Child Maltreatment, 14*(1), 17-26.

Kolb, D. L., Beutler, L. E., Davis, C. S., Crago, M., & Shanfield, S. B. (1985). Patient and therapy process variables relating to dropout and change in psychotherapy. *Psychotherapy, 22,* 702–710.

Korfmacher, J., Kitzman, H. K., & Olds, D. L. (1998). Intervention processes as conditioners of home visitation program effects. *Journal of Community Psychology, 26*, 49–64.

Korfmacher, J., Green, B., Spellman, M., & Thornburg, K. (2007). The helping relationship and program participation in early childhood home visiting. *Infant Mental Health Journal, 28*(5), 459-480.

Kotch, J B, Browne, D. C., Dufort, V., & Winsor, J. (1999). Predicting child maltreatment in the first 4 years of life from characteristics assessed in the neonatal period. *Child Abuse & Neglect*, *23*(4), 305-19.

Kotch, J., Browne, D. & Ringwalt, C. (1995). Risk of child abuse or neglect in a cohort of low-income children *Child Abuse and Neglect*, *19*, 1115-1130.

Kristman, V., Manno, M., & Côté, P. (2004). Loss to follow-up in cohort studies: how much is too much?. *European Journal of Epidemiology*, *19*(8), 751-760.

Krugman, R. D. (1993). Universal home visiting: A recommendation from the US Advisory Board on Child Abuse and Neglect. *The Future of Children*, *3*(3), 184-191.

Laflamme, L., Hasselberg, M., & Burrows, S. (2010). 20 Years of research on socioeconomic inequality and children's unintentional injuries understanding the cause-specific evidence at Hand. *International Journal of Pediatrics, 2010*, 1-23.

Lambert, D., Donahue, A., Mitchell, M., & Strauss, R. (2003). *Mental Health Outreach: Promising Practices in Rural Areas.* Rockville , MD : US Department of Health and

Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.

Landsverk, J., Carillio, T., Connelly, C. D., Ganger,W. C., Slymen, D. J., Newton, R. R., Leslie, L., & Jones, C. (2002). *Healthy Families San Diego clinical trial: Technical report*. San Diego Children's Hospital and Health Center, Child and Adolescent Services Research Center.

Lanehart, R., Rodriguez de Gil, P., Kim, E., Bellara, A., Kromrey, J., & Lee, R. (2012). Propensity score analysis and assessment of propensity score approaches using SAS procedures. *SAS Global Forum, Statistics and Data Analysis*, paper 314-2012.

Lanier, P., Jonson-Reid, M., Stahlschmidt, M. J., Drake, B., & Constantino, J. (2010). Child maltreatment and pediatric health outcomes: a longitudinal study of low-income children. *Journal of Pediatric Psychology*, *35*(5), 511-22.

LeCroy, C. & Whitaker, K. (2005). Improving the quality of home visitation: An exploratory study of difficult situations. *Child Abuse and Neglect, 29,* 1003-1013.

Li, F., Yu, Y., and Rubin, D.B. (2012). Imputing Missing Data by Fully Conditional Models: Some Cautionary Examples and Guidelines. *Duke University Department of Statistical Science Discussion Paper, 11-24*.

Little, R. J., & Rubin, D. B. (1987). *Statistical analysis with missing data* (Vol. 539). New York: Wiley.

Logsdon M. & Hutti M. (2006) Readability: an important issue impacting healthcare for women with postpartum depression. *American Journal of Maternal and Child Nursing, 31*(6), 350–355.

Loman, L. A. (2006). *Families Frequently Encountered by Child Protection Services.* Institute of Applied Research, St. Louis, MO.

Loman, A.L., Shannon, C., Sapokaite, L., & Siegel, G.L. (2009). *Minnesota Parent Support Outreach Program Evaluation: Final Report*. Institute of Applied Research, St. Louis, MO.

Love, J. M. (2002). *Making a Difference in the Lives of Infants and Toddlers and Their Families the Impacts of Early Head Start: Executive Summary*. Child Outcomes Research and Evaluation, Office of Planning, Research, and Evaluation, Administration for Children and Families and the Head Start Bureau, Administration on Children, Youth and Families, Department of Health and Human Services.

Love, S., Sanders, M., Metzler, C., Prinz, R., & Kast, E. (2013). Enhancing accessibility and engagement in evidence-based parenting programs to reduce maltreatment: Conversations with vulnerable parents. *Journal of Public Child Welfare, 7*(1), 20-38.

Luby, J. L., Barch, D. M., Belden, A., Gaffrey, M. S., Tillman, R., Babb, C., & Botteron, K. N. (2012). Maternal support in early childhood predicts larger hippocampal volumes at school age. *Proceedings of the National Academy of Sciences*, *109*(8), 2854-2859.

Lundahl, B. W., Kunz, C., Brownell, C., Tollefson, D., & Burke, B. L. (2010). A meta-analysis of motivational interviewing: twenty-five years of empirical studies. *Research on Social Work Practice*, *20*(2), 137-160.

Lynch, M., & Cicchetti, D. (1998). An ecological-transactional analysis of children and contexts: the longitudinal interplay among child maltreatment, community violence, and children's symptomatology. *Development and Psychopathology*, *10*(2), 235-57.

MacKenzie, M. J., Kotch, J. B., & Lee, L. C. (2011). Toward a cumulative ecological risk model for the etiology of child maltreatment. *Children and Youth Services Review*, *33*(9), 1638-1647.

MacLeod, J., & Nelson, G. (2000). Programs for the promotion of family wellness and the prevention of child maltreatment: a meta-analytic review. *Child Abuse & Neglect*, *24*(9), 1127-49.

MacMillan, H. L. (2010). Commentary: Child maltreatment and physical health: a call to action. *Journal of Pediatric Psychology*, *35*(5), 533-5.

MacMillan, H. L., Thomas, B. H., Jamieson, E., Walsh, C., Boyle, M. H., Shannon, H. S., & Gafni, A. (2005). Effectiveness of home visitation by public-health nurses in prevention of the recurrence of child physical abuse and neglect: a randomised controlled trial. *Lancet*, *365*(9473), 1786-93.

MacMillan, H., Wathen, C., Barlow, J., Fergusson, D., Leventhal, J., & Taussig, H. (2009). Interventions to prevent child maltreatment and associated impairment. *Lancet*, *373*(9659), 250-266.

Marcenko, M. (1996). Outcomes of a home visitation trial for pregnant and postpartum women at-risk for child placement. *Children and Youth Services Review*, *18*(3), 243-259.

Masten, A. & Coatsworth, J. (1998). The development of competence in favorable and unfavorable environments: Lessons from research on successful children. *American Psychologist, 53*, 205-220.

Mathews T.J. & MacDorman M. F. (2010). Infant mortality statistics from the 2006 period linked birth/infant death data set. *National Vital Statistics Reports, 58*(17). Hyattsville, MD: National Center for Health Statistics. 2010.

McCabe, B. K., Potash, D., Omohundro, E., & Taylor, C. R. (2012). Design and implementation of an integrated, continuous evaluation, and quality improvement system for a State-based home-visiting program. *Maternal and Child Health Journal*, *16*(7), 1385-1400.

McCrae, J. S., & Barth, R. P. (2008). Using cumulative risk to screen for mental health problems in child welfare. *Research on Social Work Practice*, *18*(2), 144-159.

McCrae, J. S., Cahalane, H., & Fusco, R. (2011). Directions for developmental screening in child welfare based on the ages and stages questionnaires. *Children and Youth Services Review*, *33*(8), 1412-1418.

McCurdy, K., & Daro, D. (2001). Parent involvement in family support programs: An integrated theory. *Family Relations, 50*(2), 113– 121.

McCurdy, K., Daro, D., Anisfeld, E., Katzev, A., Keim, A., Lecroy, C., McAfee, C., Nelson, C., Falconnier, L., McGuigan, W. M., Park, J. K., Sandy, J., & Winje, C. (2006). Understanding maternal intentions to engage in home visiting programs. *Children and Youth Services Review*, *28*(10), 1195–1212.

McGuigan, W. M., Katzev, A. R., & Pratt, C. C. (2003). Multi-level determinants of retention in a home-visiting child abuse prevention program. *Child Abuse & Neglect, 27*(4), 363-380.

McNaughton, D. B. (2008). A synthesis of qualitative home visiting research. *Public Health Nursing, 17*(6), 405–14.

Meckstroth, A., Burwick, A., & Moore, Q. (2008). *Teaching self-sufficiency : An impact and benefit- cost analysis of a home visitation and life skills education program, Findings from the rural welfare-to-work strategies*. Princeton, NJ. Mathematica Policy Institute.

Mendel, P., Meredith, L. S., Schoenbaum, M., Sherbourne, C. D., & Wells, K. B. (2008). Interventions in organizational and community context: a framework for building

evidence on dissemination and implementation in health services research. *Administration and Policy in Mental Health and Mental Health Services Research*, *35*(1-2), 21-37.

Mercy, J., & Saul, J. (2009). Creating a healthier future through early interventions for children. *The Journal of the American Medical Association*, *301*(21), 2262-2264.

Messer, S. & Wampold, B. (2002). Let's face facts: Common factors are more potent than specific therapy ingredients. *Clinical Psychology: Science and Practice, 9*(1), 21-25.

Mikton, C., & Butchart, A. (2009). Child maltreatment prevention: a systematic review of reviews. *Bulletin of the World Health Organization*, *87*(5), 353-361.

Miller, W. R. (1983). Motivational interviewing with problem drinkers. *Behavioural Psychotherapy*, *11*(02), 147-172.

Milner, J. (1986). *The Child Abuse Potential Inventory: Manual (2nd ed.).* Webster, NC: Psytec.

Milner, J. (1994). Assessing physical child abuse potential: The child abuse potential inventory. *Clinical Psychology Review, 14*(6), 547-583.

Milner, J. S., & Crouch, J. L. (1999). *Child physical abuse: Theory and research*. In R. L. Hampton, Family violence: Prevention and treatment (pp. 33-65). SAGE.

Mistry, R., Biesanz, J., Taylor, L., Burchinal, M. & Cox, M. (2004). Family income and its relation to preschool children's adjustment for families in the NICHD study of early child care. *Developmental Psychology, 40*, 727-745.

Murray, D., & Cox, J. L. (1990). Screening for depression during pregnancy with the Edinburgh Depression Scale (EDDS). *Journal of Reproductive and Infant Psychology*, *8*(2), 99-107.

Myers, J. A., & Louis, T. (2010). Regression adjustment and stratification by propensity score in treatment effect estimation. *Johns Hopkins University, Dept. of Biostatistics Working Papers*, (Working Paper 203).

Nansel, T. R., Weaver, N. L., Jacobsen, H. A., Glasheen, C., & Kreuter, M. W. (2008). Preventing unintentional pediatric injuries: a tailored intervention for parents and providers. *Health Education Research, 23*(4), 656-669.

National Research Council and Institute of Medicine, Board on Children, Youth, and Families, Commission on Behavioral and Social Sciences and Education, Committee on Integrating the Science of Early Childhood Development. J. P. Shonkoff & D. A. Phillips (Eds.) (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

National Research Council. (1993). *Understanding child abuse and neglect*. National Academy Press.

Nemeroff, C. (2004). Neurobiological consequences of childhood trauma. *Journal of Clinical Psychiatry, 65*(Supp 1), 18-21.

Newberger, C. M., & Newberger, E. (1982). Prevention of child abuse: Theory, myth, practice. *Journal of Preventive Psychiatry*, *1*(4), 443-451.

Neyman, J. (1923). On the application of probability theory to agricultural experiments: Essay on principles. Section 9. *Statistical Science, 5*, 465-480.

Nock, M. K., & Kazdin, A. E. (2005). Randomized controlled trial of a brief intervention for increasing participation in parent management training. *Journal of Consulting and Clinical Psychology*, *73*(5), 872.

Oakes, J. M., & Kaufman, J. S. (Eds.). (2006). *Methods in social epidemiology*. San Francisco. Wiley-Bass.

Olds, D. L. (2006). The Nurse-Family Partnership: An evidence-based preventive intervention. *Infant Mental Health Journal*, *27*(1), 5-25.

Olds, D.L. (2008). Preventing child maltreatment and crime with prenatal and infancy support of parents: The Nurse-Family Partnership. *Journal of Scandinavian Studies of Criminology Crime Prevention*, *9*(S1), 2-24.

Olds, D. L., Eckenrode, J., & Kitzman, H. (2005). Clarifying the impact of the Nurse-Family Partnership on child maltreatment: Response to Chaffin (2004). *Child Abuse & Neglect, 29*(3), 229-233.

Olds, D. L., Henderson, C. R., Chamberlin, R., & Tatelbaum, R. (1986). Preventing child abuse and neglect: A randomized trial of nurse home visitation. *Pediatrics*, *78*(1), 65-78.

Olds, D. L., Henderson, C. R., Kitzman, H. J., Eckenrode, J. J., Cole, R. E., & Tatelbaum, R. C. (1999). Prenatal and infancy home visitation by nurses: Recent findings. *Future of Children, 9*(1), 44–65.

Olds, D. L., Henderson, C. R., Kitzman, H., & Cole, R. (1995). Effects of prenatal and infancy nurse home visitation on surveillance of child maltreatment. *Pediatrics, 95,* 365–372.

Olds, D. L., Kitzman, H., Cole, R., & Robinson, J. (1997). Theoretical foundations of a program of home visitation for pregnant women and parents of young children. *Journal of Community Psychology*, *25*(1), 9-25.

Olds, D. L., Kitzman, H., Cole, R., Robinson, J., Sidora, K., Luckey, D.W., Henderson, C.R.J., Hanks, C., Bondy, J., & Holmberg, J. (2004). Effects of nurse home-visiting on maternal

life course and child development: age 6 follow-up results of a randomized trial. *Pediatrics, 114*(6), 1550-1559.

Olds, D. L., & Korfmacher, J. (1998). Maternal psychological characteristics as influences on home visitation contact. *Journal of Community Psychology*, *26*(1), 23-36.

Olivieri, I., Bova, S. M., Urgesi, C., Ariaudo, G., Perotto, E., Fazzi, E., & Orcesi, S. (2012). Outcome of extremely low birth weight infants: What's new in the third millennium? Neuropsychological profiles at four years. *Early Human Development*, *88*(4), 241-250.

Olson, S., Bates, J., Sandy, J., & Schilling, E. Early developmental precursors of impulsive and inattentive behavior: From infancy to middle childhood. *Journal of Child Psychology and Psychiatry, 43*, 435-447.

Olson, M. E., Diekema, D., Elliott, B., & Renier, C. M. (2010). Impact of income and income inequality on infant health outcomes in the United States. *Pediatrics*, *126*(6), 1165–73.

Palusci, V. J. (2011). Risk factors and services for child maltreatment among infants and young children. *Children and Youth Services Review, 33*(8), 1374-1382.

Parke, R., & Collmer, C. (1975). Child abuse: An interdisciplinary analysis. In E. Hetherington (Ed.), *Review of Child Development Research*. Oxford, England: University of Chicago Press.

Parrish, J. W., & Gessner, B. D. (2010). Infant maltreatment-related mortality in Alaska: correcting the count and using birth certificates to predict mortality. *Child Abuse & Neglect, 34*(12), 951-958.

Paulsell, D., Avellar, S., Martin, E. S., & Del Grosso, P. (2010). *Home visiting evidence of effectiveness review: Executive summary November 2010*. Washington, DC: Office of

Planning, Research and Evaluation, Administration for Children and Families, US DHHS.

Pecora, P., Whittaker, J. K., & Maluccio, A. N. (2000). *The child welfare challenge*. New York: Aldine Transaction.

Perez, C. M., & Widom, C. S. (1994). Childhood victimization and long-term intellectual and academic outcomes. *Child Abuse & Neglect*, *18*(8), 617-33.

Perry, B.D. (2000). *The neuroarcheology of childhood maltreatment: The neurodevelopmental costs of adverse childhood events.* Retrieved from http://www.childtrauma.org/ctamaterials/Neuroarcheology.asp.

Phillips, J., Charles, M., Sharpe, L., & Matthey, S. (2005). Validation of the subscales of the Edinburgh Postnatal Depression Scale in a sample of women with unsettled infants. *Archives Of Women's Mental Health*, *8*(2), 89–95.

Pollack, S. (2005, Jan). *Emotion processing and neurobehavioral correlates*. Presented at Child Neglect Consortium Annual Meeting, Washington, DC: National Institutes of Health.

Prinz, R. J., Sanders, M. R., Shapiro, C. J., Whitaker, D. J., & Lutzker, J. R. (2009). Population-based prevention of child maltreatment: the U.S. Triple P system population trial. *Prevention Science*, *10*(1), 1-12.

Proctor, E. K., Landsverk, J., Aarons, G., Chambers, D., Glisson, C., & Mittman, B. (2009). Implementation research in mental health services: An emerging science with conceptual, methodological, and training challenges. *Administration and Policy in Mental Health and Mental Health Services Research*, *36*(1), 24-34.

Putnam-Hornstein, E., & Needell, B. (2011). Predictors of child protective service contact between birth and age five: An examination of California's 2002 birth cohort. *Children and Youth Services Review, 33*(8), 1337-1344.

Raver, C. & Leadbeater, B. (1999). Mothering under pressure: Environmental, child, and dyadic correlates of maternal self-efficacy among low-income women. *Journal of Family Psychology, 13,* 523-534.

Reynolds, A., Mathieson, L. & Topitzes, J. (2009). Do early childhood interventions prevent child maltreatment? A review of research. *Child Maltreatment, 14*(2), 182-206.

Ringeisen, H., Casanueva, C., Urato, M., & Cross, T. (2008). Special health care needs among children in the child welfare system. *Pediatrics, 122*(1), e232-41.

Roberts, I., Kramer, M. S., & Suissa, S. (1996). Does home visiting prevent childhood injury? A systematic review of randomised controlled trials. *British Medical Journal*, *312,* 29-33.

Robinson, J., Emde, R. (2004). Mental health moderators of Early Head Start on parenting and child development: Maternal depression and relationship attitudes. *Parenting: Science and Practice, 4*, 73-97.

Rohrbach, S. (1993). Healthy from the start. Home care for newborns. *Caring, 12*(12), 13-5.

Rosenbaum, P. R., & Rubin, D. R. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55.

Rosenbaum, P. R. & Rubin, D. B. (1985) Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician, 39*, 33-38.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688-701.

Rubin, D.B. (1987). *Multiple imputation for nonresponse in surveys*. New York: J. Wiley & Sons.

Rubin, D. B. (2010). Propensity score methods. *American Journal of Ophthalmology, 149*(1), 7-9.

Rubin, D.B. & Schenker, N. (1986). Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *Journal of the American Statistical Association, 81*(394), 366-374.

Rutter, M. (2007). Resilience, competence, and coping. *Child Abuse and Neglect, 31*, 205-209.

Salt, A., & Redshaw, M. (2006). Neurodevelopmental follow-up after preterm birth: follow up after two years. *Early Human Development, 82*(3), 185-197.

Sameroff, A. J. (1998). Environmental risk factors in infancy. *Pediatrics,102*(Supplement E1), 1287-1292.

Sameroff, A. J. (2000). Developmental systems and psychopathology. *Development and Psychopathology, 12*, 297-312.

Sameroff, A. J., Seifer, R., Barocas, R., Zax, M., & Greenspan, S. (1987). Intelligence quotient scores of 4-year-old children: Social-environmental risk factors. *Pediatrics*, *79*(3), 343-350.

Sanders, M. R. (1999). Triple P-Positive Parenting Program: Towards an empirically validated multilevel parenting and family support strategy for the prevention of behavior and emotional problems in children. *Clinical Child and Family Psychology Review*, *2*(2), 71-90.

Scheffer, J. (2002). Dealing with missing data. *Research Letters in the Informational and Mathematical Sciences, 3*(1), 153-160.

Schnitzer, P. G., Slusher, P. L., Kruse, R. L., & Tarleton, M. M. (2011). Identification of ICD codes suggestive of child maltreatment. *Child Abuse & Neglect, 35*(1), 3-17.

Scott, D., Tonmyr, L., Fraser, J., Walker, S., & McKenzie, K. (2009). The utility and challenges of using ICD codes in child maltreatment research: A review of existing literature. *Child Abuse & Neglect, 33*(11), 791-808.

Scribano, P. (2010). Prevention strategies in pediatrics. *Current Opinion in Pediatrics, 22*(5), 616-620.

Sedlak, A., Mettenburg, J., Basena, M., Petta, I., McPherson, K., Greene, A., & Li, S. (2010). *Fourth National Incidence Study of Child Abuse and Neglect (NIS – 4): Report to Congress, Executive Summary*. Washington, DC.

Segre, L., Stasik, S., O'Hara, M., & Arndt, S. (2010). Listening visits: An evaluation of the effectiveness and acceptability of a home-based depression treatment. *Psychotherapy Research, 20*(6), 712-721

Sheeder, J., Kabir, K., & Stafford, B. (2009). Screening for postpartum depression at well-child visits: is once enough during the first 6 months of life? *Pediatrics, 123*(6), e982-e988.

Shonkoff, J.P. & Garner, A.S. (2012). American Academy of Pediatrics Technical Report. The Lifelong Effects of Early Childhood Adversity and Toxic Stress. *Pediatrics, 129*(1): e232-e244

Shonkoff, J. P., Boyce, W. T., & McEwen, B. S. (2009). Neuroscience, molecular biology, and the childhood roots of health disparities: building a new framework for health promotion and disease prevention. *Journal of the American Medical Association, 301*(21), 2252-2259.

Sidebotham, P. (2001). An ecological approach to child abuse: a creative use of scientific models in research and practice. *Child Abuse Review*, *10*(2), 97-112.

Silovsky, J., Bard, D., Chaffin, M., Hecht, D., Burris, L., Owora, A., Beasley, L., Doughty, D., & Lutzker, J. (2011). Prevention of child maltreatment in high-risk rural families: A randomized clinical trial with child welfare outcomes. *Children and Youth Services Review, 33,* 1435-1444.

Singer, L. T., Salvator, A., Guo, S., Collin, M., Lilien, L., & Baley, J. (1999). Maternal psychological distress and parenting stress after the birth of a very low-birth-weight infant. *JAMA: The Journal of the American Medical Association*, *281*(9), 799-805.

Singh, G. K., & Kogan, M. D. (2007). Persistent socioeconomic disparities in infant, neonatal, and postneonatal mortality rates in the United States, 1969-2001. *Pediatrics*, *119*(4), e928–39.

Slack, K. S., Holl, J. L., Lee, B. J., McDaniel, M., Altenbernd, L., & Stevens, A. B. (2003). Child protective intervention in the context of welfare reform: The effects of work and welfare on maltreatment reports. *Journal of Policy Analysis and Management*, *22*(4), 517-536.

Squires, J., Bricker, D., & Potter, L. (1997). Revision of a parent-completed developmental screening tool: Ages and Stages Questionnaires. *Journal of Pediatric Psychology*, *22*(3), 313-328.

Squires, J., Potter, L., & Bricker, D. D. (1999). *The ASQ user's guide*. Baltimore, MD: Paul H. Brookes.

Stagner, M. W., & Lansing, J. (2009). Progress toward a prevention perspective. *The Future of Children*, *19*(2), 19-38.

Steele, B. F., & Pollack, G. (1974). A psychiatric study of parents who abuse their children and infants. In C. H. Kempe (Ed.), *The battered child*. Chicago: University of Chicago Press.

Stith, S. M., Liu, T., Davies, L. C., Boykin, E. L., Alder, M. C., Harris, J. M., ... & Dees, J. E. M. E. G. (2009). Risk factors in child maltreatment: A meta-analytic review of the literature. *Aggression and Violent Behavior*, *14*(1), 13-29.

Strathearn, L., Gray, P., O'Callaghan, M. & Wood, D. (2001). Childhood neglect and cognitive development in extremely low birth weight infants: A prospective study. *Pediatrics, 108,* 142-151.

Straus, M., Hamby, S. L., Finkelhor, D., Moore, D. W., & Runyan, D. (1998). Identification of child maltreatment with the Parent-Child Conflict Tactics Scales: development and psychometric data for a national sample of American parents. *Child Abuse & Neglect*, *22*(4), 249-70.

Sweet, M., & Appelbaum, M. I. (2004). Is home visiting an effective strategy? A meta-analytic review of home visiting programs for families with young children. *Child Development*, *75*(5), 1435-56.

Taube, D. O., & Elwork, A. (1990). Researching the effects of confidentiality law on patients' self-disclosures. *Professional Psychology, Research and Practice*, *21*(1), 72-5.

Taylor, C., Guterman, N. B., Lee, S. J., & Rathouz, P. J. (2009). Intimate partner violence, maternal stress, nativity, and risk for maternal maltreatment of young children. *American Journal of Public Health*, *99*(1), 175-83.

Teicher, M., Dumont, N., Ito, Y., Vaituzis, C., Giedd, J., & Andersen, S. (2004). Childhood neglect is associated with reduced corpus callosum area. *Biological Psychiatry, 56*, 80-85.

Thoemmes, F. J. & Kim, E. S. (2011). A systematic review of propensity score methods in the social sciences. *Multivariate Behavioral Research, 46(1)***,** 90-118.

Thomas, D., Leichet, C., Hughes, C., Madigan, A., & Dowell, K. (2003). *Emerging practices in the prevention of child abuse and neglect*. Washington, DC. US Department of Health and Human Services.

Thompson, L., Kropenske, V., Heinicke, C. M., Gomby, D. S., & Halfon, N. (2001). *Home visiting: A service strategy to deliver Proposition 10 Results*. UCLA Center for Healthier Children, Families, and Communities.

Thyer, B. A. (2008). The quest for evidence-based practice: we are all positivists! *Research on Social Work Practice*, *18*(4), 339-345.

Tzeng, O. C., Jackson, J. W., & Karlson, H. C. (1991). Theories of child abuse and neglect. Differential perspectives, summaries, and evaluations. Westport, Connecticut: Praeger.

U.S. Department of Health and Human Services US DHHS. (1999). *Mental Health: A Report of the Surgeon General—Executive Summary*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health.

U.S. Department of Health and Human Services, Administration on Children and Families, Administration on Children, Youth and Families, Children's Bureau. (2010). *Child Maltreatment 2009.* Washington, DC: Available from http://www.acf.hhs.gov/programs/cb/stats_research/index.htm#can.

U.S. Department of Health and Human Services US DHHS. (2012). Office of Disease Prevention and Health Promotion. Healthy People 2020. Washington, DC.

U.S. Dept. of Health & Human Services Health Resources & Services Administration, Bureau of Primary Health Care. (1996). *Models that work: compendium of innovative primary health care programs for underserved and vulnerable populations 1996*. Bethesda, MD: Author.

Wagner, M. M., & Clayton, S. L. (1999). The Parents as Teachers program: Results from two demonstrations. *Future of Children, 9*(1), 91–115.

Walsh, C., MacMillan, H., Jamieson, E., (2002). The relationship between parental psychiatric disorder and child physical and sexual abuse: Findings from the Ontario Health Supplement. *Child Abuse & Neglect*, *26*, 11-22.

Wang, C. & Holton, J. (2007). *Total estimated cost of child abuse and neglect in the United States statistical evidence.* Chicago (IL): Prevent Child Abuse America (PCAA).

Welsh, B., Sullivan, C., & Olds, D. (2010). When early crime prevention goes to scale: A new look at the evidence. *Prevention Science*, *11*(2), 115-125.

Whitaker, D. J., Lutzker, J. R., & Shelley, G. (2005). Child maltreatment prevention priorities at the Centers for Disease Control and Prevention. *Child Maltreatment*, *10*(3), 245-59.

Widom, C. (2000). Understanding the consequences of childhood victimization. In R. Reece (Ed.), *Treatment of child abuse: Common ground for mental health, medical, and legal practitioners*. Baltimore and London: Johns Hopkins University Press.

Widom, C., Raphael, K. G., & DuMont, K. (2004). The case for prospective longitudinal studies in child maltreatment research: Commentary on Dube, Williamson, Thompson, Felitti, and Anda (2004). *Child Abuse & Neglect*, *28*(7), 715-22.

Widom, C., Weiler, B. L., & Cottler, L. B. (1999). Childhood victimization and drug abuse: a comparison of prospective and retrospective findings. *Journal of Consulting and Clinical Psychology*, *67*(6), 867-80.

Windham, A., Rosenberg, L., Fuddy, L., McFarlane, E., Sia, C., & Duggan, A. (2004). Risk of mother-reported child abuse in the first 3 years of life. *Child Abuse & Neglect*, *28*(6), 645-667.

Wu, S. S., Ma, C. X., Carter, R. L., Ariet, M., Feaver, E. A., Resnick, M. B., & Roth, J. (2004). Risk factors for infant maltreatment: a population-based study. *Child Abuse & Neglect*, *28*(12), 1253-1264.

Wulczyn, F. (2009). Epidemiological perspectives on maltreatment prevention. *The Future of Children, (19)* 2, 39-66.

Ying, G. S., & Liu, C. (2006). Statistical analysis of clustered data using SAS system. In *Proceedings of the North East SAS Users Group (NESUG) Conference* (pp. 17-20).

Zelkowitz, P., Bardin, C., & Papageorgiou, A. (2007). Anxiety affects the relationship between parents and their very low birth weight infants. *Infant Mental Health Journal*, *28*(3), 296-313.

Zielinski, D. S., & Bradshaw, C. P. (2006). Ecological influences on the sequelae of child maltreatment: a review of the literature. *Child Maltreatment*, *11*(1), 49-62.

Zielinski, D., Eckenrode, J., & Olds, D. (2009). Nurse home visitation and the prevention of child maltreatment: Impact on the timing of officials reports. *Development and Psychopathology, 21*, 441-453.

Zigler. E., Pfannenstiel, J & Sietz, V. (2008). The Parents as Teachers program and school success: A replication and extension. *Journal of Primary Prevention, 29*, 103-120.

Appendix A: Empirical Evidence Assessing Home Visiting Effectiveness in Maltreatment Prevention

Reference	**Intervention/Setting**	**Sample Inclusion/Exclusion**	**CA/N Assessment**	**Findings**
Dumont, et al., 2008	Healthy Families New York RCT based on HFA model Screen expectant parents and those with infants under 3 months then assess risk with FSC Visits by paraprofessionals until age 5	1,173 families in three sites, 34% white, 45% AA, 18% Latina, 54% first-time mothers, 29% welfare Data from baseline, year 1 and year 2 Prevention subgroup of first-time mothers and psychologically vulnerable (MH and mastery) subgroup	Self-report (PC-CTS) Substantiated CPS reports	No program effects for prevalence of any self-reported subscales At year 1, intervention group reported engaging in fewer acts of serious physical abuse, minor physical aggression, and psychological aggression At year 2, fewer acts of physical abuse and neglect No significant differences in prevalence or frequency of substantiated reports
Duggan, et al., 2007	Healthy Families Alaska Experimental study of 6 programs families randomized 3-5 years of home visiting by 34 paraprofessionals	325 families interviewed at baseline and child age 2 Families who screen positive and >25 on FSC 23% Alaskan Native, 54% Caucasian, 57% below poverty	Observed (HOME) and self-reported parenting (PC-CTS) behaviors Hospitalization for trauma or inadequate care (ACSCs)(maternal interview and primary care records) Maternal relinquishment (follow-up interviews) Substantiated CPS reports	Little evidence of effectiveness in preventing child abuse although high risk sample No differences in reports, relinquishment, ACSCs, ED use, self-reported discipline, neglect, use of community services, etc. Less likely to provide poor quality HOME and less use of milder physical and psychological discipline No evidence of moderation of outcomes or dose effect
Minkovitz, et al., 2007	Healthy Steps for Young Children Universal, practice-based intervention	Follow-up at age 5 of 3,165 children Socioeconomically advantaged sample,	Telephone interview of maternal self-report of response to child misbehavior	Intervention group less likely to report slapping or hitting with object ($p<.001$) but similar rates of use of harsh discipline

	including 6 home visits in the three years 3 year controlled trial with 15 sites	65% white, 33% Medicaid		
MacMillan, et al. 2005	RCT of program in Canada to prevent recurrent maltreatment Visitation by a public health nurse	163 families referred to local CPS agencies, child younger than 13 yo, 90% low-income	Subsequent physical abuse or neglect of any child in the family based on CPS reports 3 years after randomization Abusive parenting (CAPI) Home environment (HOME Inventory)	Recurrence of maltreatment did not differ between groups. No difference on CAPI or HOME between groups or change over time
Fergusson, et al., 2005	Early Start RCT in New Zealand Family support workers are bachelor's level nurses or social workers	443 families, 90% welfare dependent, 27% Maori Referrals by Plunket community nurses for families screened for more than 2 HSP risk factors	Self-reported parenting (PC-CTS) Parent report of contact with Child, Youth, and Family Service	Significantly lower rate of severe physical assault (OR=.35, p<.01) No difference in rates of agency contact
Duggan, et al., 2004	Hawaii Healthy Start Program Effectiveness study taken to scale Trained paraprofessionals Visits for 3-5 years weekly, biweekly, monthly, then quarterly	643 families randomly assigned to intervention or control English speaking HSP or hospital staff screen for risk by maternal medical record or >25 on Family Stress Checklist 34% Native Hawaiian, 10% Caucasian, 63% below poverty	Observed (HOME) and self-reported parenting (PC-CTS) behaviors Hospitalization for trauma or inadequate care (ACSCs)(maternal interview and primary care records) Maternal relinquishment (follow-up interviews) Substantiated CPS reports	HSP and control groups were similar on most measures of maltreatment. Less likely to use corporal/verbal punishment (OR=.59, p=.01) with agency-specific effect Maternal reported less neglect (OR=.72, p=.02) No family subgroup differences Favorable and unfavorable dose effects

Fraser, et al., 2000	RCT in Queensland, Australia Interdisciplinary model, parents visited by child health nurse, social worker, and parent aides coordinated by community pediatrician Minimum of 18 visits per family until age 1	Women recruited from a hospital in the postnatal period with at least one major risk factor OR three minor risk factors 90 Intervention, 91 Comparison, 44% primiparas, 6% Aboriginal, 75% low income	Child abuse risk (CAPI) at baseline, 7 months, and 18 months Observation/interview of the quality of the home environment (HOME Inventory) at 6 weeks and 12 months	Statistically significant difference in CAPI score at 7 months (11% vs. 30% elevated), effect maintained at follow-up Significantly different HOME scores at 6 weeks but not at 12 months
Kitzman, et al., 1997	Nurse Family Partnership Memphis 4-arm RCT: transportation to prenatal care, screening and referral, nurse home visitation during pregnancy, visitation during and until 2nd birthday	1129 women less than 29 weeks, primiparas, at least two of (unmarried, no HS education, or unemployed) 92% AA, 85% below poverty Average of 7 prenatal and 26 postnatal visits	Interviewed at 28 and 36th week of pregnancy and 6, 12, 24th month Observation of properties of home environment (HOME Inventory) Medical records of ED visits, injuries, and ingestions	Nurse-visited children had fewer healthcare encounters (p=.05) and hospitalization (p<.001)for injuries and ingestions Nurse-visited homes were more conducive to children's development (p=.003)
Olds, Chamberlain, & Tatelbaum, 1986 Follow-up: Olds, et al., 1997	Nurse Family Partnership 4-arm RCT (screening and referral, screening and free transportation for services, home visiting during pregnancy, home	400 pregnant women with no previous live births At least one risk factor (<19 yo, unmarried, Medicaid status or no private insurance) 11% AA, 59% low SES Follow-up study of children at age 15 (81%	State DSS reviewed records at age 2 for "indicated" reports CPS records from all states resided in until age 15 to ascertain total number of substantiated reports involving mother as perpetrator	At 2-year follow-up, a trend (4% vs. 19%, p=.07) identified only for highest risk group, poor and unmarried teens Mothers visited during pregnancy and the first 2 years were identified as perpetrators in fewer reports at fifteen year follow-up (control incidence=.54, treatment=.29, p<.01)

	visiting through age 2) Semirural Elmira, NY recruited from clinic from 1978 to 1980 Average of 9 prenatal and 23 home visits through age 2	of original)		Effect was greater for unmarried and low-SES mothers (control incidence=.53, treatment=.11, p<.01)
Marcenko & Spence, 1994	Interdisciplinary home visitation model for pregnant women at-risk for out-of-home placement Services provided by a peer home visitor with a social worker and nurse Services from first prenatal visit to first birthday, minimum of every two weeks	125 experimental, 100 control, urban setting with primarily AA (94%) women, 84% welfare, average of 4 pregnancies Recruited from outpatient clinic with at least one of (substance abuse, homelessness, domestic violence, psychiatric illness, incarceration, HIV, lack of social support)	Maternal report of CPS involvement, repeated at follow-up Observation of quality stimulation available to the infant (HOME Inventory) Follow-up after 10 months and 16 months from baseline	At first follow-up, more experimental group women had children in placement (32% vs. 19%), but same at 12 months Experimental group more likely to have family placement than foster care and more likely to reunify No difference in overall quality of home environment Control group scored significantly higher only on organization of home environment subscale
Barth, 1991	California Child Parent Enrichment Project Random assignment to paraprofessional (parenting consultant) home visits for six months using task-centered approach	Pregnant women referred by public health, education, or social service professionals 97 intervention and 94 control 44% primiparas, 45% white, 31% Latino, 17% AA, 40% AFDC	Mother's well-being (CAPI) Self-report caregiving problems (child removed from care, neighbor has cared for child) CPS intake calls, substantiated, and unsubstantiated reports average of three years	CAPI scores not significantly different at posttest Follow-up reports of child abuse were similar for both groups

			later	
Siegel, et al., 1980	Randomly assigned at delivery to home visits by paraprofessional infant care worker 9 home visits during first 3 months	321 women in 3rd trimester at public prenatal clinic	Interviews and observations (Attachment Inventory) at 4 and 12 months, hospital and CPS records at 12 months	No statistically significant effects on maternal attachment, reports of child abuse and neglect, or health care utilization

Appendix B: Review Articles and Meta-Analyses Reviewing Effectiveness of Home Visiting for Child Outcomes

Review Article	Type/Articles Included	Conclusions
Paulsell, Avellar, Martin, & Del Grosso (2010). Home visiting evidence of effectiveness review	Thorough and transparent review of home visiting literature Database search, website search, and call for studies	The review identified several gaps in the existing research literature on home visiting models that limit its usefulness for matching program models to community needs. First, research evidence of program effectiveness is limited. As noted earlier, many models do not have high- or moderate-quality studies of their effectiveness; thus, policymakers and program administrators cannot determine whether those models are effective. Other models have only a few high- or moderate-quality studies, indicating that additional research on those models may be needed. Second, more evidence is needed about the effectiveness of home visiting models for different types of families with a range of characteristics. Overall, the studies had fairly diverse study samples in terms of race/ethnicity and income. However, sample sizes in these studies are not typically large enough to allow for analysis of findings separately by subgroup. Moreover, HomVEE found little or no research on the effectiveness of home visiting program models for families from American Indian tribes, immigrant families that have diverse cultural backgrounds or may not speak English as a first language, or military families.
Macmillan, et al., (2009). Interventions to prevent maltreatment and associated impairment.	Review article of strategies to prevent maltreatment	Despite the promotion of a broad range of early childhood home-visiting programmes, most of these have not been shown to reduce physical abuse and neglect when assessed using RCTs. Some systematic reviews, especially those including meta-analyses, have concluded that early childhood home visitation is effective in preventing child abuse and neglect without taking into account the variability across programmes. Such general statements obscure important differences in design and methods, including outcomes, across studies. Two programmes, the Nurse–Family Partnership developed in the USA and the Early Start programme in New Zealand have, however,

		shown significant benefits.
Howard & Brooks-Gunn (2009). The role of home-visiting programs in preventing child abuse and neglect.	*Future of Children* review article of home visiting as a strategy to prevent child abuse and neglect	Overall, researchers have found little evidence that home-visiting programs directly prevent child abuse and neglect. But home visits can impart positive benefits to families by way of influencing maternal parenting practices, the quality of the child's home environment, and children's development. Programs have their greatest benefits for low-income, first-time adolescent mothers. If home-visiting programs are to have their maximum impact, service providers must follow carefully the guidelines mandated by the respective programs, use professional staff whose credentials are consistent with program goals, intervene prenatally with at-risk populations, and carry out the programs with fidelity to their theoretical models.
Gomby. (2005). Home visitation in 2005: Outcomes for children and parents	Committee for Economic Development Investing in Kids working paper Review paper	Home visiting services can produce the results that prepare children for school, but they do not always do so in practice. And, benefits are often small. When averaged across program models, sites, and families, results for most outcomes are about .1 or .2 of a standard deviation in size, an effect size that is considered small in human services. Effects are most consistent for outcomes related to parenting. Home visiting programs do not generate consistent benefits in child development or in improving the course of mothers' lives. Families in which children have obvious risk factors (e.g., they are biologically at-risk, developmentally delayed, or they already have behavior problems) appear to benefit most. Some studies also suggest that the highest-risk mothers (e.g., low income teen mothers; mothers with poor coping skills, low IQs, and mental health problems) may benefit most. For every outcome, as many as half of the studies and programs demonstrate extremely small or no benefits at all. But, for every outcome, a few programs or program sites demonstrate larger benefits, and it is those more positive results which have driven the expansion of home visiting programs and which

		illustrate the potential of home visiting.
Biluhka, et al. (2005). The effectiveness of early childhood home visitation in preventing violence, a systematic review.	Systematic review of articles published before July 2001 Reviewed interventions effectiveness and economic efficiency	Available studies provide strong evidence that early childhood home visitation programs are effective in preventing child maltreatment, reducing reported maltreatment by approximately 39%. Programs delivered by professional visitors (nurses or mental health workers) seem to yield greater effects than those delivered by paraprofessionals. For paraprofessional visitors, effects are mixed, and beneficial effects are generally found in programs of longer duration.
Sweet & Appelbaum (2004). Is home visiting an effective strategy? A meta-analytic review of home visiting programs for families with young children.	Meta-analysis of 60 programs	In general, children in families who were enrolled in home visiting programs fared better than did control group children. Within the set of child outcomes, three of the five average effect sizes were significantly greater than zero. Only child abuse and parent stress as an indicator of potential for child abuse did not yield an average effect size significantly greater than zero.
MacLeod & Nelson (2000). Programs for the promotion of family wellness and the prevention of child maltreatment: A meta-analytic review	Meta-analysis of 23 home-visiting programs	Home visitation effect size equals .41, findings suggest that home visitation programs should last more than 6 months and provide more than 12 home visits; those lasting less than 6 months and providing 12 or fewer visits do not appear to be very effective in preventing child maltreatment. Home visitation interventions achieved higher effect sizes with participants of mixed SES than participants with low SES.
Kendrick, et al., (2000). Does home visiting improve parenting and	Systematic review and meta-analysis of 34	Our review of the effectiveness of home visiting programmes suggests they are effective in increasing the quality of the home environment as measured by HOME scores, and that the majority of studies using other outcome

quality of the home environment? A systematic review and meta-analysis.	studies	measures also indicated significant improvements in a variety of measures of parenting. While the majority of the studies we reviewed focused on families living in socioeconomic deprivation, it should not be assumed that "poor parenting" is the preserve of such families; or that inequalities in terms of material resources do not need addressing.
Gomby, Culross, & Behrman (1999). Home visiting: Recent program evaluations – analysis and recommendations.	*Future of Children* review article on evaluations of home visiting evaluations	Evaluation findings are "sobering". In most of the studies described, programs struggled to enroll, engage, and retain families. When program benefits were demonstrated, they usually accrued only to a subset of the families originally enrolled in the programs, they rarely occurred for all of a program's goals, and the benefits were often quite modest in magnitude. Authors recommend a dedicated effort, led by the field, to improve the quality and implementation of existing home visiting services, and a more modest view of the potential of the broad array of home visiting programs.
Roberts, Kramer, & Suissa (1996). Does home visiting prevent childhood injury? A systematic review of randomized controlled trials.	Systematic review of 11 RCTs that examined child injury or abuse	Home visiting programmes have the potential to reduce significantly the rates of childhood injury. The problem of differential surveillance for child abuse between intervention and control groups precludes the use of reported abuse as a valid outcome measure in controlled trials of home visiting
Olds & Kitzman (1993). Review of research on home visiting for pregnant women and parents of young children.	*Future of Children* review article on evaluations of home visiting evaluations	In summary, none of the six trials that sought to use home visiting to prevent child abuse and neglect demonstrated overall decreases in maltreatment as evidenced by state CPS records. Three, however, did demonstrate differences for at least some study participants which are suggestive of benefits, either in decreasing abuse or neglect, improving parenting, or decreasing use of medical services often associated with abuse and neglect.